(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 251 446 B1

(12) # EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: 28.12.94

(51) Int. Cl.5: C12N 15/00, C12N 9/54, C12N 1/00

(21) Application number: 87303761.8

(22) Date of filing: 28.04.87

(54) **Non-human Carbonyl hydrolase mutants, DNA sequences and vectors encoding same and hosts transformed with said vectors.**

(30) Priority: 30.04.86 US 858594
06.04.87 US 35652

(43) Date of publication of application:
07.01.88 Bulletin 88/01

(45) Publication of the grant of the patent:
28.12.94 Bulletin 94/52

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(56) References cited:
EP-A- 0 130 756
WO-A-87/04461
WO-A-87/05050

ABSTRACTS OF THE 190TH AMERICAN CHEMICAL SOCIETY NATIONAL MEETING, vol. 190,1985, page 23, no. 47; R.R. BOTT et al.: "Protein engineering of subtilisin"

(73) Proprietor: GENENCOR INTERNATIONAL, INC.
180 Kimball Way
South San Francisco, CA 94080 (US)

(72) Inventor: Wells, James Allen
64 Otay Avenue
San Mateo
CA 94403 (US)
Inventor: Cunningham, Brian C.
24 Olive Avenue
Piedmont
CA 94611 (US)
Inventor: Caldwell, Robert Mark
1828 Broadway
No.101
San Francisco
Ca 94109 (US)
Inventor: Bott, Richard Ray
3032 Hillside drive
Burlingame
CA 94010 (US)

EP 0 251 446 B1

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 10, part A, 1986, page271, no. E101, SYMPOSIUM ON PROTEASES IN BIOLOGICAL CONTROL AND BIOTECHNOLOGY,15th ANNUAL UCLA, MEETING ON MOLECULAR AND CELLULAR BIOLOGY, Los Angeles, CA.,9th-15th February 1986; P. BRYAN et al.: "Protein engineering of subtilisin-proteases of enhanced stability"

WORLD BIOTECH. REPORT, vol. 2, 1985, pages 51-59, Online Publications, Pinner,GB; R. BOTT: "Modeling & crystallographic analysis of site-specific mutants of subtilisin"

JOURNAL OF CELLULAR BIOCHEMISTRY SUPPL., vol. 0, no. 11, part C, 1987, page 200, no. N024, New York, US; D.A. ESTELL et al.: "Tailoring enzymatic properties through multiple mutations"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 84, March 1987, pages 1219-1223, Washington, D.C., US; J.A. WELLS et al.: "Designing substrate specifity by protein engineering of electrostatic interactions"

BIOCHEMISTRY, vol. 26, no. 8, April 1987, pages 2077-2082, American Chemical Society, Washington, D.C., US; M.W. PANTOLIANO et al.: "Protein engineering of subtilisin BPN': enhanced stabilization through the introduction of two cysteines to form a disulfide bond"

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCE USA, vol. 83, June 1986, pages 3743-3745, Washington, D.C., US; P. BRYAN et al.: "Site-directed mutagenesis and the role of the oxyanion hole in subtilisin"

NATURE, vol. 318, 28th November 1985, pages 375-376, London, GB; P.G. THOMAS etal.: "Tailoring the pH dependence of enzyme catalysis using proteinengineering"

JOURNAL OF BACTERIOLOGY, vol. 158, no. 2, May 1984, pages 411-418, American Society for Microbiology, Washington, D.C., US; M.L. STAHL et al.: "Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro-derived deletion mutation"

Inventor: **Estell, David Aaron**
250 Diablo Avenue
Mountan View
CA 94043 (US)
Inventor: **Power, Scott Douglas**
732 Olive Court
San Bruno
CA 94066 (US)

(74) Representative: **Armitage, Ian Michael et al**
MEWBURN ELLIS
York House
23 Kingsway
London WC2B 6HP (GB)

NUCLEIC ACIDS RESEARCH, vol. 11, no. 22, November 1983, pages 7911-7925, IRL Press Ltd, Cambridge, GB; J.A. WELLS et al.: "Cloning, sequencing, and secretion of Bacillus amyloliquefaciens subtilisin in Bacillus subtilis"	

Description

The recent development of various in vitro techniques to manipulate the DNA sequences encoding naturally-occuring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the speculation that such techniques can be used to modify enzymes to improve some functional property in a predictable way. Ulmer, K.M. (1983) Science 219, 666-671. The only working example disclosed therein is the substitution of a single amino acid within the active site of tyrosyl-tRNA synthetase (Cys35→Ser) which lead to a reduction in enzymatic activity. See Winter, G., et al. (1982) Nature 299, 756-758; and Wilkinson, A.J., et al. (1983) Biochemistry 22, 3581-3586 (Cys35→Gly mutation also resulted in decreased activity).

When the same t-RNA synthetase was modified by substituting a different amino acid residue within the active site with two different amino acids, one of the mutants (Thr51→Ala) reportedly demonstrated a predicted moderate increase in kcat/Km whereas a second mutant (Thr51→Pro) demonstrated a massive increase in kcat/Km which could not be explained with certainty. Wilkinson, A.H., et al. (1984) Nature 307, 187-188.

Another reported example of a single substitution of an amino acid residue is the substitution of cysteine for isoleucine at the third residue of T4 lysozyme. Perry, L.J., et al. (1984) Science 226, 555-557. The resultant mutant lysozyme was mildly oxidized to form a disulfide bond between the new cysteine residue at position 3 and the native cysteine at position 97. This crosslinked mutant was initially described by the author as being enzymatically identical to, but more thermally stable than, the wild type enzyme. However, in a "Note Added in Proof", the author indicated that the enhanced stability observed was probably due to a chemical modification of cysteine at residue 54 since the mutant lysozyme with a free thiol at Cys54 has a thermal stability identical to the wild type lysozyme.

Similarly, a modified dihydrofolate reductase from E.coli has been reported to be modified by similar methods to introduce a cysteine which could be cross linked with a naturally-occurring cysteine in the reductase. Villafranca, D.E., et al. (1983) Science 222, 782-788. The author indicates that this mutant is fully reactive in the reduced state but has significantly diminished activity in the oxidized state. In addition, two other substitutions of specific amino acid residues are reported which resulted in mutants which had diminished or no activity.

EPO Publication No. 0130756 discloses the substitution of specific residues within B. amyloliquefaciens subtilisin with specific amino acids. Thus, Met222 has been substituted with all 19 other amino acids, Gly166 with 9 different amino acids and Gly169 with Ala and Ser.

As set forth below, several laboratories have also reported the use of site directed mutagensis to produce the mutation of more than one amino acid residue within a polypeptide.

The amino-terminal region of the signal peptide of the prolipoprotein of the E. coli outer membrane was stated to be altered by the substitution or deletion of residues 2 and 3 to produce a charge change in that region of the polypeptide. Inoyye, S., et al. (1982) Proc. Nat. Acad. Sci. USA 79, 3438-3441. The same laboratory also reported the substitution and deletion of amino acid redisues 9 and 14 to determine the effects of such substitution on the hydrophobic region of the same signal sequence. Inouye, S., et al. (1984) J. Biol. Chem. 259, 3729-3733.

Double mutants in the active site of tyrosyl-t-RNA synthetase have also been reported. Carter, P.J., et al. (1984) Cell 38, 835-840. In this report, the improved affinity of the previously described Thr51→Pro mutant for ATP was probed by producing a second mutation in the active site of the enzyme. One of the double mutants, Gly35/Pro51, reportedly demonstrated an unexpected result in that it bound ATP in the transition state better than was expected from the two single mutants. Moreover, the author warns, at least for one double mutant, that it is not readily predictable how one substitution alters the effect caused by the other substitution and that care must be taken in interpreting such substitutions.

A mutant is disclosed in U.S. Patent No. 4,532,207, wherein a polyarginine tail was attached to the C-terminal residue of β-urogastrone by modifying the DNA sequence encoding the polypeptide. As disclosed, the polyarginine tail changed the electrophoretic mobility of the urogastrone-polyaginine hybrid permiting selective purification. The polyarginine was subsequently removed, according to the patentee, by a polyarginine specific exopeptidase to produce the purified urogastrone. Properly construed, this reference discloses hybrid polypeptides which do not constitute mutant polypeptides containing the substitution, insertion or deletion of one or more amino acids of a naturally occurring polypeptide.

Single and double mutants of rat pancreatic trypsin have also been reported. Craik, C.S., et al. (1985) Science 228, 291-297. As reported, glycine residues at positions 216 and 226 were replaced with alanine residues to produce three trypsin mutants (two single mutants and one double mutant). In the case of the single mutants, the authors stated expectation was to observe a differential effect on Km. They instead

reported a change in specificity (kcat/Km) which was primarily the result of a decrease in kcat. In contrast, the double mutant reportedly demonstrated a differential increase in Km for lysyl and arginyl substrates as compared to wild type trypsin but had virtually no catalytic activity.

The references discussed above are provided solely for their disclosure prior to the filing date of the instant case, and nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or priority based on earlier filed applications.

Based on the above references, however, it is apparent that the modification of the amino acid sequence of wild type enzymes often results in the decrease or destruction of biological activity.

Accordingly, it is an object herein to provide carbonyl hydrolase mutants which have at least one property which is different from the same property of the carbonyl hydrolase precursor from which the amino acid of said mutant is derived.

It is a further object to provide mutant DNA sequences encoding such carbonyl hydrolase mutants as well as expression vectors containing such mutant DNA sequences.

Still further, another object of the present invention is to provide host cells transformed with such vectors as well as host cells which are capable of expressing such mutants either intracellularly or extracellularly.

Summary of the Invention

The invention includes carbonyl hydrolase mutants, preferably having at least one property which is substantially different from the same property of the precursor non-human carbonyl hydrolase from which the amino acid sequence of the mutant is derived. These properties include oxidative stability, substrate, specificity catalytic activity, thermal stability, alkaline stability, pH activity profile and resistance to proteolytic degradation. The precursor carbonyl hydrolase may be naturally occurring carbonyl hydrolases or recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is derived by the substitution, deletion or insertion of one or more amino acids of the precursor carbonyl hydrolase amino acid sequence.

The invention also includes mutant DNA sequences encoding such carbonyl hydrolase mutants. Further the invention includes expression vectors containing such mutant DNA sequences as well as host cells transformed with such vectors which are capable of expressing said carbonyl hydrolase mutants.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence of the coding strand, correlated with the amino acid sequence of B. amyloliquefaciens subtilisin gene. Promoter (p) ribosome binding site (rbs) and termination (term) regions of the DNA sequence as well as sequences encoding the presequence (PRE) putative prosequence (PRO) and mature form (MAT) of the hydrolase are also shown.

Figure 2 is a schematic diagram showing the substrate binding cleft of subtilisin together with substrate.

Figure 3 is a stereo view of the S-1 binding subsite of B. amyloliquefaciens subtilisin showing a lysine P-1 substrate bound in the site in two different ways. Figure 3A shows Lysine P-1 substrate bound to form a salt bridge with a Glu at position 156. Figure 3B shows Lysine P-1 substrate bound to form a salt bridge with Glu at position 166.

Figure 4 is a schematic diagram of the active site of subtilisin Asp32, His64 and Ser221.

Figures 5A and 5B depict the amino acid sequence of subtilisin obtained from various sources. The residues directly beneath each residue of B. amyloliquefaciens subtilisin are equivalent residues which (1) can be mutated in a similar manner to that described for B. amyloliquefaciens subtilisin, or (2) can be used as a replacement amino acid residue in B. amyloliquefaciens subtilisin. Figure 5C depicts conserved residues of B. amyloliquefaciens subtilisin when compared to other subtilisin sequences.

Figures 6A and 6B depict the inactivation of the mutants Met222L and Met222Q when exposed to various organic oxidants.

Figure 7 depicts the ultraviolet spectrum of Met222F subtilisin and the difference spectrum generated after inactivation by diperdodecanoic acid (DPDA).

Figure 8 shows the pattern of cyanogen bromide digests of untreated and DPDA oxidized subtilisin Met222F on high resolution SDS-pyridine peptide gels.

Figure 9 depicts a map of the cyanogen bromide fragments of Fig. 8 and their alignment with the sequence of subtilisin Met222F.

Figure 10 depicts the construction of mutations between codons 45 and 50 of B. amyloliquefaciens subtilisin.

Figure 11 depicts the construction of mutations between codons 122 and 127 of B. amyloliquefaciens subtilisin.

Figure 12 depicts the effect of DPDA on the activity of subtilisin mutants at positions 50 and 124 in subtilisin Met222F.

Figure 13 depicts the construction of mutations at codon 166 of B. amyloliquefaciens subtilisin.

Figure 14 depicts the effect of hydrophobicity of the P-1 substrate side-chain on the kinetic parameters of wild-type B. amyloliquefaciens subtilisin.

Figure 15 depicts the effect of position 166 side-chain substitutions on P-I substrate specificity. Figure 15A shows position 166 mutant subtilisins containing non-branched alkyl and aromatic side-chain substitutions arranged in order of increasing molecular volume. Figure 15B shows a series of mutant enzymes progressing through β- and γ-branched aliphatic side chain substitutions of increasing molecular volume.

Figure 16 depicts the effect of position 166 side-chain volumn on log kcat/Km for various P-1 substrates.

Figure 17 shows the substrate specificity differences between Ile166 and wild-type (Gly166) B. amyloliquefaciens subtilisin against a series of alphatic and aromatic substrates. Each bar represents the difference in log kcat/Km for Ile166 minus wild-type (Gly166) subtilisin.

Figure 18 depicts the construction of mutations at codon 169 of B. amyloliquefaciens subtilisin.

Figure 19 depicts the construction of mutations at codon 104 of B. amyloliquefaciens subtilisin.

Figure 20 depicts the construction of mutations at codon 152 B. amyloliquefaciens subtilisin.

Figure 21 depicts the construction of single mutations at codon 156 and double mutations at codons 156 and 166 of B. amyloliquefaciens subtilisin.

Figure 22 depicts the construction of mutations at codon 217 for B. amyloliquefaciens subtilisin.

Figure 23 depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 23A depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 24 depicts the kcat/Km versus pH profile for mutations at codon 222 in B. amyloliquefaciens subtilisin.

Figure 25 depicts the constructing mutants at codons 94, 95 and 96.

Figures 26 and 27 depict substrate specificity of various wild type and mutant subtilisins for different substrates.

Figures 28 A, B, C and D depict the effect of charge in the P-1 binding sites due to substitutions at codon 156 and 166.

Figures 29 A and B are a stereoview of the P-1 binding site of subtilisin BPN' showing a lysine P-1 substrate bound in the site in two ways. In 29A, Lysine P-1 substrate is built to form a salt bridge with a Glu at codon 156. In 29B, Lysine P-1 substrate is built to form a salt bridge with Glu at codon 166.

Figure 30 demonstrates residual enzyme activity versus temperature curves for purified wild-type (Panel A), C22/C87 (Panel B) and C24/C87 (Panel C).

Figure 31 depicts the strategy for producing point mutations in the subtilisin coding sequence by misincorporation of α-thioldeoxynucleotide triphosphates.

Figure 32 depicts the autolytic stability of purified wild type and mutant subtilisins 170E, 107V, 213R and 107V/213R at alkaline pH.

Figure 33 depicts the autolytic stability of purified wild type and mutant subtilisins V50, F50 and F50/V107/R213 at alkaline pH.

Figure 34 depicts the strategy for constructing plasmids containing random cassette mutagenesis over residues 197 through 228.

Figure 35 depicts the oligodeoxynucleotides used for random cassette mutagenesis over residues 197 through 228.

Figure 36 depicts the construction of mutants at codon 204.

Figure 37 depicts the oligodeoxynucleotides used for synthesizing mutants at codon 204.

Detailed Description

The inventors have discovered that various single and multiple in vitro mutations involving the substitution, deletion or insertion of one or more amino acids within a non-human carbonyl hydrolase amino acid sequence can confer advantageous properties to such mutants when compared to the non-mutated carbonyl hydrolase.

Specifically, B. amyloliquefaciens subtilisin, an alkaline bacterial protease, has been mutated by modifying the DNA encoding the subtilisin to encode the substitution of one or more amino acids at various amino acid residues within the mature form of the subtilisin molecule. These in vitro mutant subtilisins have at least one property which is different when compared to the same property of the precursor subtilisin. These modified properties fall into several categories including: oxidative stability, substrate specificity, thermal stability, alkaline stability, catalytic activity, pH activity profile, resistance to proteolytic degradation, Km, kcat and Km/kcat ratio.

Carbonyl hydrolases are enzymes which hydrolyze compounds containing

bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include α-aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exoproteases.

"Recombinant carbonyl hydrolase" refers to a carbonyl hydrolase in which the DNA sequence encoding the naturally occurring carbonyl hydrolase is modified to produce a mutant DNA sequence which encodes the substitution, insertion or deletion of one or more amino acids in the carbonyl hydrolase amino acid sequence. Suitable modification methods are disclosed herein and in EPO Publication No. 0130756 published January 9, 1985.

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted by various bacterial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus is aspartate-histidineserine. In the chymotrypsin related proteases the relative order, however is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

"Recombinant subtilisin" refers to a subtilisin in which the DNA sequence encoding the subtilisin is modified to produce a mutant DNA sequence which encodes the substitution, deletion or insertion of one or more amino acids in the naturally occurring subtilisin amino acid sequence. Suitable methods to produce such modification include those disclosed herein and in EPO Publication No. 0130756. For example, the subtilisin multiple mutant herein containing the substitution of methionine at amino acid residues 50, 124 and 222 with phenylalanine, isoleucine and glutamine, respectively, can be considered to be derived from the recombinant subtilisin containing the substitution of glutamine at residue 222 (Q222) disclosed in EPO Publication No. 0130756. The multiple mutant thus is produced by the substitution of phenylalanine for methionine at residue 50 and isoleucine for methionine at residue 124 in the Q222 recombinant subtilisin.

"Carbonyl hydrolases" and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as E. coli or pseudomonas and gram positive bacteria such as micrococcus or bacillus. Examples of eucaryotic organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as S. cerevisiae, fungi such as Aspergillus sp., and non-human mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless exhibit the same or similar type of biological activity. Thus, non-human carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and non-human eucaryotic sources.

A "carbonyl hydrolase mutant" has an amino acid sequence which is derived from the amino acid sequence of a non-human "precursor carbonyl hydrolase". The precursor carbonyl hydrolases include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is "derived" from the precursor hydrolase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the amino acid sequence of the precursor carbonyl hydrolase rathern than manipulation of the precursor carbonyl hydrolase per se. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO Publication No. 0130756.

Specific residues of B. amyloliquefaciens subtilisin are identified for substitution, insertion or deletion. These amino acid position numbers refer to those assigned to the B. amyloliquefaciens subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor carbonyl hydrolases containing amino acid residues which are "equivalent" to the particular identified residues in B. amyloliquefaciens subtilisin.

A residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of B. amyloliquefaciens subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analagous to a specific residue or portion of that residue in B. amyloliquefaciens subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor carbonyl hydrolase is directly comparted to the B. amyloliquefaciens subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 5C). After aligning the conserved residues, allowing for necessary insertions and deletions in order to maintain alignment (i.e., avoiding the elimination of conserved residues through arbitrary deletion and insertion), the residues equivalent to particular amino acids in the primary sequence of B. amyloliquefaciens subtilisin are defined. Alignment of conserved residues preferably should conserve 100% of such residues. However, alignment of greater than 75% or as little as 50% of conserved residues is also adequate to define equivalent residues. Conservation of the catalytic triad, Asp32/His64/Ser221 should be maintained.

For example, in Figure 5A the amino acid sequence of subtilisin from B. amyloliquefaciens B. subtilisin var. I168 and B. lichenformis (carlsbergensis) are aligned to provide the maximum amount of homology between amino acid sequences. A comparison of these sequences shows that there are a number of conserved residues contained in each sequence. These residues are identified in Fig. 5C.

These conserved residues thus may be used to define the corresponding equivalent amino acid residues of B. amyloliquefaciens subtilisin in other carbonyl hydrolases such as thermitase derived from Thermoactinomyces. These two particular sequences are aligned in Fig. 5B to produce the maximum homology of conserved residues. As can be seen there are a number of insertions and deletions in the thermitase sequence as compared to B. amyloliquefaciens subtilisin. Thus, in thermitase the equivalent amino acid of Tyr217 in B. amyloliquefaciens subtilisin is the particular lysine shown beneath Tyr217.

In Fig. 5A, the equivalent amino acid at position 217 in B. amyloliquefaciens subtilisin is Tyr. Likewise, in B. subtilis subtilisin position 217 is also occupied by Tyr but in B. licheniformis position 217 is occupied by Leu.

Thus, these particular residues in thermitase, and subtilisin from B. subtilisin and B. licheniformis may be substituted by a different amino acid to produce a mutant carbonyl hydrolase since they are equivalent in primary structure to Tyr217 in B. amyloliquefaciens subtilisin. Equivalent amino acids of course are not limited to those for Tyr217 but extend to any residue which is equivalent to a residue in B. amyloliquefaciens whether such residues are conserved or not.

Equivalent residues homologous at the level of tertiary structure for a precursor carbonyl hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the precursor carbonyl hydrolase and B. amyloliquefaciens subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of atomic coordinates of non-hydrogen protein atoms of the carbonyl hydrolase in question to the B. amyloliquefaciens subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest resolution available.

$$R\ factor = \frac{\sum_h |Fo(h)| - |Fc(h)|}{\sum_h |Fo(h)|}$$

Equivalent residues which are functionally analogous to a specific residue of B. amyloliquefaciens subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the B. amyloliquefaciens subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie with 0.13nm of the corresponding side chain atoms of B. amyloliquefaciens subtilisin. The three dimensional structures would be aligned as outlined above.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a mutant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally occurring sequence. The carbonyl hydrolase mutants of the present invention include the mature forms of carbonyl hydrolase mutants as well as the pro- and prepro-forms of such hydrolase mutants. The prepro-forms are the preferred construction since this facilitates the expression, secretion and maturation of the carbonyl hydrolase mutants.

"Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in EPO Publication No. 0130756 to render them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing subtilisin is the Bacillus strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in EPO Publicatin No. 0130756 and further described by Yang, M.Y., et al. (1984) J. Bacteriol. 160, 15-21. Other host cells for expressing subtilisin include Bacillus subtilis I168 (EPO Publication No. 0130756).

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the carbonyl hydrolase mutants or expressing the desired carbonyl hydrolase mutant. In the case of vectors which encode the pre or prepro form of the carbonyl hydrolase mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor carbonyl hydrolase may be obtained in accord with the general methods described herein in EPO publication No. 0130756.

Once the carbonyl hydrolase gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-occurring precursor carbonyl hydrolase. Such modifications include the production of recombinant carbonyl hydrolases as disclosed in EPO

Publication No. 0130756 and the production of carbonyl hydrolase mutants described herein.

The carbonyl hydrolase mutants of the present invention may be generated by site specific mutagenesis (Smith, M. (1985) Ann. Rev. Genet. 423; Zoeller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500), cassette mutagenesis (EPO Publication No. 0130756) or random mutagenesis (Shortle, D., et al. (1985) Genetics, 110, 539; Shortle, D., et al. (1986) Proteins: Structure, Function and Genetics, 1, 81; Shortle, D. (1986) J. Cell. Biochem, 30, 281; Alber, T., et al. (1985) Proc. Natl. Acad. of Sci., 82, 747; Matsumura, M., et al. (1985) J. Biochem., 260, 15298; Liao, H., et al. (1986) Proc. Natl. Acad. of Sci., 83 576) of the cloned precursor carbonyl hydrolase. Cassette mutagenesis and the random mutagenesis method disclosed herein are preferred.

The mutant carbonyl hydrolases expressed upon transformation of suitable hosts are screened for enzymes exhibiting one or more properties which are substantially different from the properties of the precursor carbonyl hydrolases, e.g., changes in substrate specificity, oxidative stability, thermal stability, alkaline stability, resistance to proteolytic degradation, pH-activity profiles and the like.

A change in substrate specificity is defined as a difference between the kcat/Km ratio of the precursor carbonyl hydrolase and that of the hydrolase mutant. The kcat/Km ratio is a measure of catalytic efficienty. Carbonyl hydrolase mutants with increased or diminished kcat/Km ratios are described in the examples. Generally, the objective will be to secure a mutant having a greater (numerically large) kcat/Km ratio for a given substrate, thereby enabling the use of the enzyme to more efficiently act on a target substrate. A substantial change in kcat/Km ratio is preferably at least 2-fold increase or decrease. However, smaller increases or decreases in the ratio (e.g., at least 1.5-fold) are also considered substantial. An increase in kcat/Km ratio for one substrate may be accompanied by a reduction in kcat/Km ratio for another substrate. This is a shift in substrate specificity, and mutants exhibiting such shifts have utility where the precursor hydrolase is undesirable, e.g. to prevent undesired hydrolysis of a particular substrate in an admixture of substrates. Km and kcat are measured in accord with known procedures, as described in EPO Publication No. 0130756 or as described herein.

Oxidative stability is measured either by known procedures or by the methods described hereinafter. A substantial change in oxidative stability is evidenced by at least about 50% increase or decrease (preferably decrease) in the rate of loss of enzyme activity when exposed to various oxidizing conditions. Such oxidizing conditions are exposure to the organic oxidant diperdodecanoic acid (DPDA) under the conditions described in the examples.

Alkaline stability is measured either by known procedures or by the methods described herein. A substantial change in alkaline stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half life of the enzymatic activity of a mutant when compared to the precursor carbonyl hydrolase. In the case of subtilisins, alkaline stability was measured as a function of autoproteolytic degradation of subtilisin at alkaline pH, e.g. for example, 0.1M sodium phosphate, pH 12 at 25° or 30°C.

Thermal stability is measured either by known procedures or by the methods described herein. A substantial change in thermal stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half-life of the catalytic activity of a mutant when exposed to a relatively high temperature and neutral pH as compared to the precursor carbonyl hydrolase. In the case of subtilisins, thermal stability is measured by the autoproteolytic degradation of subtilisin at elevated temperatures and neutral pH, e.g., for example 2mM calcium chloride, 50mM MOPS pH 7.0 at 59°C.

The inventors have produced mutant subtilisins containing the substitution of the amino acid residues of B. amyloliquefaciens subtilisin shown in Table I. The wild type amino acid sequence and DNA sequence of B. amyloliquefaciens subtilisin is shown in Fig. 1.

TABLE I

Residue	Replacement Amino Acid
Tyr21	F A
Thr22	C
Ser24	C
Asp32	Q S
Ser33	A T
Asp36	A G
Gly46	V
Ala48	E V R
Ser49	C L
Met50	C F V
Asn77	D
Ser87	C
Lys94	C
Val95	C
Leu96	D
Tyr104	A C D E F G H I K L M N P Q R S T V W
Ile107	V
Gly110	C R
Met124	I L
Asn155	A D H Q T
Glu156	Q S
Gly166	C E I L M P S T W Y
Gly169	C D E F H I K L M N P Q R T V W Y
Lys170	E R
Tyr171	F
Pro172	E Q
Phe189	A C D E G H I K L M N P Q R S T V W Y
Asp197	R A
Met199	I
Ser204	C R L P
Lys213	R T
Tyr217	A C D E F G H I K L M N P Q R S T V W
Ser221	A C

The different amino acids substituted are represented in Table I by the following single letter designations:

Amino acid or residue thereof	3-letter symbol	1-letter symbol
Alanine	Ala	A
Glutamate	Glu	E
Glutamine	Gln	Q
Aspartate	Asp	D
Asparagine	Asn	N
Leucine	Leu	L
Glycine	Gly	G
Lysine	Lys	K
Serine	Ser	S
Valine	Val	V
Arginine	Arg	R
Threonine	Thr	T
Proline	Pro	P
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Cysteine	Cys	C
Tryptophan	Trp	W
Histidine	His	H

Except where otherwise indicated by context, wild-type amino acids are represented by the above three-letter symbols and replaced amino acids by the above single-letter symbols. Thus, if the methionine at residue 50 in B. amyloliquefaciens subtilisin is replaced by phenylalanine, this mutation (mutant) may be designated Met50F or F50. Similar designations are used for multiple mutants.

In addition to the amino acids used to replace the residues disclosed in Table I, other replacements of amino acids at these residues are expected to produce mutant subtilisins having useful properties. These residues and replacement amino acids are shown in Table II.

TABLE II

Residue	Replacement Amino Acid(s)
Tyr-21	L
Thr22	K
Ser24	A
Asp32	
Ser33	G
Gly46	
Ala48	
Ser49	
Met50	L K I V
Asn77	D
Ser87	N
Lys94	R Q
Val95	L I
Tyr104	
Met124	K A
Ala152	C L I T M
Asn155	
Glu156	A T M L Y
Gly166	
Gly169	
Tyr171	K R E Q
Pro172	D N
Phe189	
Tyr217	
Ser221	
Met222	

Each of the mutant subtilisins in Table I contain the replacement of a single residue of the B. amyloliquefaciens amino acid sequence. These particular residues were chosen to probe the influence of such substitutions on various properties of B. amyloliquefacien subtilisin.

Thus, the inventors have identified Met124 and Met222 as important residues which if substituted with another amino acid produce a mutant subtilisin with enhanced oxidative stability. For Met124, Leu and Ile are preferred replacement amino acids. Preferred amino acids for replacement of Met222 are disclosed in EPO Publication No. 0130756.

Various other specific residues have also been identified as being important with regard to substrate specificity. These residues include Tyr104, Ala152, Glu156, Gly166, Gly169, Phe189 and Tyr217 for which mutants containing the various replacement amino acids presented in Table I have already been made, as well as other residues presented below for which mutants have yet to be made.

The identification of these residues, including those yet to be mutated, is based on the inventors' high resolution crystal structure of B. amyloliquefaciens subtilisin to 1.8 A (see Table III), their experience with in vitro mutagenesis of subtilisin and the literature on subtilisin. This work and the x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L., et al. (1976) J. Biol. Chem. 251, 1097-1103), has helped in identifying an extended peptide binding cleft in subtilisin. This substrate binding cleft together with substrate is schematically diagramemed in Fig. 2, according to the nomenclature of Schechter, I., et al. (1967) Biochem Bio. Res. Commun. 27, 157. The scissile bond in the substrate is identified by an arrow. The P and P' designations refer to the amino acids which are positioned respectively toward the amino or carboxy terminus relative to the scissle bond. The S and S' designations refer to subsites in the substrate binding cleft of subtilisin which interact with the corresponding substrate amino acid residues.

Atomic Coordinates for the
Apoenzyme Form of B, Amyloliquefaciens
Subtilisin to 1.8AResolution

1	ALA	N	19.434	53.195	-21.756
1	ALA	C	18.731	50.925	-21.324
1	ALA	CB	21.099	51.910	-21.103
2	GLN	CA	17.219	49.000	-21.434
2	GLN	O	18.765	47.165	-21.691
2	GLN	CG	15.920	47.905	-21.927
2	GLN	OE1	13.023	40.612	-22.067
3	SER	N	17.477	47.205	-19.052
3	SER	C	16.735	44.910	-19.430
3	SER	CB	18.500	45.030	-18.069
4	VAL	N	16.991	43.646	-19.725
4	VAL	C	16.129	41.934	-18.290
4	VAL	CB	16.800	41.622	-20.022
4	VAL	CG2	16.037	42.266	-22.106
5	PRO	CA	15.304	41.415	-16.027
5	PRO	O	14.005	39.263	-17.146
5	PRO	CG	13.041	43.215	-15.921
6	TYR	N	16.363	39.240	-15.407
6	TYR	C	15.359	36.975	-15.520
6	TYR	CB	17.024	37.323	-14.034
6	TYR	CD1	18.437	35.452	-16.346
6	TYR	CE1	18.535	34.070	-16.653
6	TYR	CZ	18.222	33.154	-15.628
7	GLY	N	14.464	37.362	-14.630
7	GLY	C	12.400	36.535	-15.670
8	VAL	N	12.441	37.529	-18.541
8	VAL	C	12.363	36.433	-18.735
8	VAL	CB	11.765	38.900	-18.567
8	VAL	CG2	10.991	39.919	-17.733
9	SER	CA	14.419	35.342	-19.562
9	SER	O	14.112	33.014	-19.801
9	SER	OG	16.162	36.747	-20.350
10	GLN	CA	13.964	32.636	-16.076
10	GLN	O	12.705	30.642	-17.413
10	GLN	CG	14.295	31.617	-14.500
10	GLN	OE1	16.556	33.060	-12.746
11	ILE	N	13.625	32.575	-17.670
11	ILE	C	10.209	31.792	-19.605
11	ILE	CB	9.132	32.669	-17.475
11	ILE	CG2	9.162	32.655	-15.941
12	LYS	N	11.272	32.105	-20.277
12	LYS	C	10.456	33.006	-22.522
12	LYS	CB	11.257	30.646	-22.216
12	LYS	CD	12.563	28.517	-22.159
12	LYS	NZ	14.476	27.688	-28.035
13	ALA	CA	9.325	35.190	-22.631
13	ALA	O	9.330	35.004	-24.981
14	PRO	N	11.332	35.950	-23.093
14	PRO	C	11.706	35.957	-26.317
14	PRO	CB	13.462	36.588	-24.692
14	PRO	CD	12.201	35.936	-22.750
15	ALA	CA	13.379	33.450	-27.367
15	ALA	O	10.000	33.710	-29.270
16	LEU	N	9.003	34.130	-27.240
16	LEU	C	7.912	35.925	-28.521
16	LEU	CB	6.746	34.623	-26.690
16	LEU	CD1	5.001	33.234	-27.089
17	HIS	N	8.665	36.820	-27.922
17	HIS	C	9.510	37.901	-29.090
17	HIS	CB	9.700	39.100	-27.652
17	HIS	ND1	9.030	39.007	-25.272
17	HIS	CE1	9.226	39.914	-26.144
18	SER	N	10.463	37.033	-30.022
1	ALA	CA	19.011	51.774	-21.965
1	ALA	O	10.376	51.197	-20.175
2	GLN	N	18.260	49.006	-22.041
2	GLN	C	17.075	47.706	-20.992
2	GLN	CB	16.125	48.760	-22.449
2	GLN	CD	13.912	47.762	-22.930
2	GLN	NE2	14.115	46.917	-23.926
3	SER	CA	17.950	45.060	-19.437
3	SER	O	15.500	45.352	-19.229
3	SER	OG	17.682	46.210	-17.069
4	VAL	CA	15.946	42.619	-19.639
4	VAL	O	17.123	41.170	-18.006
4	VAL	CG1	14.076	40.972	-20.761
5	PRO	N	15.239	42.106	-17.331
5	PRO	C	15.501	39.905	-16.249
5	PRO	CB	14.150	41.000	-15.263
5	PRO	CD	14.064	42.906	-17.417
6	TYR	CA	16.620	37.803	-15.715
6	TYR	O	15.224	35.943	-16.235
6	TYR	CG	18.021	35.047	-15.055
6	TYR	CD2	17.696	34.900	-14.071
6	TYR	CE2	17.015	33.539	-14.379
6	TYR	OH	18.312	31.030	-15.996
7	GLY	CA	13.211	36.648	-14.376
7	GLY	O	11.747	35.470	-15.003
8	VAL	CA	11.777	37.523	-17.036
8	VAL	O	11.639	35.716	-19.470
8	VAL	CG1	11.186	38.093	-19.943
9	SER	N	13.661	36.310	-18.775
9	SER	C	14.100	33.920	-18.965
9	SER	CB	15.926	35.632	-19.505
10	GLN	N	14.115	33.007	-17.662
10	GLN	C	12.607	31.007	-17.277
10	GLN	CB	14.125	32.005	-15.418
10	GLN	CD	14.406	31.911	-13.147
10	GLN	NE2	14.552	30.960	-12.251
11	ILE	CA	18.373	31.984	-18.102
11	ILE	O	9.173	31.333	-20.100
11	ILE	CG1	9.066	34.117	-18.049
11	ILE	CD1	7.500	34.640	-17.923
12	LYS	CA	11.300	32.119	-21.722
12	LYS	O	10.170	32.703	-23.606
12	LYS	CG	12.203	29.038	-21.423
12	LYS	CE	13.023	27.467	-21.106
13	ALA	N	10.109	34.130	-21.991
13	ALA	C	10.026	35.716	-23.063
13	ALA	CB	8.005	36.195	-21.565
14	PRO	CA	11.905	36.430	-25.120
14	PRO	O	11.770	36.047	-27.445
14	PRO	CG	13.320	36.970	-23.221
15	ALA	N	11.560	36.236	-26.129
15	ALA	C	10.002	33.795	-28.032
15	ALA	CB	11.552	31.069	-27.062
16	LEU	CA	7.901	34.950	-27.020
16	LEU	O	7.362	36.126	-29.500
16	LEU	CG	5.700	33.665	-26.522
16	LEU	CD2	6.696	32.207	-26.203
17	HIS	CA	8.090	38.351	-28.530
17	HIS	O	9.107	38.622	-30.056
17	HIS	CG	9.105	39.200	-26.262
17	HIS	CD2	8.000	38.926	-25.694
17	HIS	NE2	8.079	39.320	-24.301
18	SER	CA	11.109	36.739	-31.322

30	ARG	C	30.190	30.123	-32.393
30	ARG	CD	32.311	35.709	-31.172
30	GLN	N	9.000	35.605	-31.063
30	GLN	C	9.162	36.131	-33.303
30	GLN	CB	9.221	33.069	-32.200
30	GLN	CD	6.023	31.707	-31.101
30	GLN	OE1	9.302	30.052	-30.296
30	GLY	CA	6.369	30.307	-32.000
30	GLY	O	6.263	39.276	-32.215
31	TYR	CA	6.310	37.031	-29.763
31	TYR	O	5.622	30.876	-27.706
31	TYR	CG	2.973	35.706	-30.709
31	TYR	CD2	3.630	36.706	-31.307
31	TYR	CE2	3.193	34.261	-32.508
31	TYR	OH	2.301	30.261	-34.290
32	THR	CA	6.262	60.927	-27.120
32	THR	O	3.207	61.725	-25.325
32	THR	OG1	6.319	62.457	-25.597
33	GLY	N	2.039	60.205	-26.693
33	GLY	C	-0.107	61.631	-26.310
34	SER	N	-0.023	61.967	-27.371
34	SER	C	-2.303	62.626	-27.064
34	SER	CB	-0.734	63.12C	-29.320
35	ASN	N	-3.050	63.602	-27.913
35	ASN	C	-5.015	62.075	-26.205
35	ASN	CB	-5.105	63.227	-20.703
35	ASN	OD1	-6.005	63.767	-31.003
36	VAL	N	-6.377	62.669	-25.292
36	VAL	C	-6.702	62.052	-22.997
36	VAL	CB	-3.716	60.903	-23.021
36	VAL	CG2	-3.090	39.576	-25.018
37	LYS	CA	-6.133	63.526	-21.175
37	LYS	O	-6.605	61.073	-19.413
37	LYS	CG	-0.066	66.975	-22.690
37	LYS	CE	-10.306	60.697	-23.137
38	VAL	N	-6.019	63.662	-19.203
38	VAL	C	-6.790	63.959	-16.020
38	VAL	CB	-2.026	62.666	-17.032
38	VAL	CG2	-2.667	61.005	-19.173
39	ALA	CA	-5.767	66.330	-16.630
39	ALA	O	-6.066	62.065	-13.106
40	VAL	N	-6.057	65.033	-13.072
40	VAL	C	-3.990	65.609	-10.601
40	VAL	CB	-2.006	65.010	-12.169
40	VAL	CG2	-2.033	65.236	-13.307
41	ILE	CA	-5.320	66.066	-8.679
41	ILE	O	-3.025	63.915	-6.097
41	ILE	CG1	-7.200	63.707	-9.709
41	ILE	CD1	-0.617	62.096	-9.717
42	ASP	CA	-2.066	66.667	-6.255
42	ASP	O	-6.307	60.610	-8.302
42	ASP	CG	-0.603	60.702	-6.270
42	ASP	OD2	-0.001	60.620	-9.330
43	SER	CA	-1.095	60.037	-6.001
43	SER	O	-1.706	32.106	-9.063
43	SER	OG	0.935	30.025	-6.776
44	GLY	CA	-2.255	31.725	-0.305
44	GLY	O	-0.166	30.031	-0.703
45	ILE	CA	0.200	32.630	-10.995
45	ILE	O	-0.327	36.635	-11.766
45	ILE	CG1	-0.330	39.210	-12.097
45	ILE	CD1	-0.962	60.605	-13.626
46	ASP	CA	2.309	39.610	-11.232
30	ARG	O	30.569	30.312	-33.036
30	ARG	OE	33.321	36.650	-30.390
30	GLN	CA	0.902	36.962	-32.070
30	GLN	O	6.207	39.072	-36.219
30	GLN	CG	9.070	32.002	-31.023
30	GLN	OE1	0.710	31.033	-31.666
30	GLY	N	9.805	37.223	-32.087
30	GLY	C	5.101	30.602	-31.000
31	TYR	N	5.202	37.001	-30.761
31	TYR	C	6.070	30.052	-28.025
31	TYR	CB	3.690	36.631	-29.663
31	TYR	CD1	1.705	36.332	-31.230
31	TYR	CE1	1.306	33.707	-32.666
31	TYR	CZ	2.003	34.755	-33.067
32	THR	N	5.902	39.600	-30.200
32	THR	C	5.091	60.022	-26.366
32	THR	CB	5.333	61.750	-27.011
32	THR	CG2	6.676	61.323	-20.220
33	GLY	CA	0.009	60.600	-25.562
33	GLY	O	-1.013	62.095	-25.930
34	SER	CA	-0.097	62.957	-20.012
34	SER	O	-2.013	61.000	-28.160
34	SER	OG	0.363	63.032	-29.720
35	ASN	CA	-6.510	63.607	-27.093
35	ASN	O	-6.233	62.660	-26.190
35	ASN	CG	-6.960	66.170	-20.035
35	ASN	ND2	-6.767	65.661	-29.096
36	VAL	CA	-6.676	61.670	-24.143
36	VAL	O	-3.050	63.619	-22.600
36	VAL	CG1	-6.160	39.002	-22.960
37	LYS	N	-3.910	62.015	-22.061
37	LYS	C	-3.015	62.072	-19.061
37	LYS	CB	-7.090	60.001	-21.169
37	LYS	CD	-9.321	65.302	-22.020
37	LYS	NZ	-9.606	66.253	-24.264
38	VAL	CA	-6.657	62.950	-17.097
38	VAL	O	-6.200	65.095	-16.017
38	VAL	CG1	-2.666	62.105	-16.509
39	ALA	N	-5.606	63.527	-15.013
39	ALA	C	-6.790	66.010	-13.053
39	ALA	CB	-7.172	66.107	-16.101
40	VAL	CA	-3.166	66.062	-11.010
40	VAL	O	-6.105	66.640	-10.070
40	VAL	CG1	-0.906	65.081	-10.098
41	ILE	N	-6.516	66.510	-0.077
41	ILE	C	-6.566	64.933	-7.060
41	ILE	CB	-6.657	63.770	-0.501
41	ILE	CG2	-7.270	66.030	-7.225
42	ASP	N	-6.066	66.193	-7.227
42	ASP	C	-3.071	67.009	-9.705
42	ASP	CB	-3.695	66.329	-7.092
42	ASP	OD1	0.094	66.902	-6.076
43	SER	N	-1.031	60.912	-9.306
43	SER	C	-1.902	30.976	-9.000
43	SER	CB	-0.623	69.022	-3.030
44	GLY	N	-2.170	30.769	-7.006
44	GLY	C	-1.030	31.660	-9.057
45	ILE	N	-0.905	32.631	-10.302
45	ILE	C	0.960	33.010	-11.263
45	ILE	CB	-0.062	31.696	-12.367
45	ILE	CG2	1.160	31.761	-13.902
46	ASP	N	1.016	34.253	-10.971
46	ASP	C	2.201	39.936	-12.702

36	ASP	O	3.004	55.471	-13.579
36	ASP	CG	6.339	57.099	-10.006
36	ASP	OD2	3.660	57.277	-10.263
37	SER	CA	3.103	57.221	-30.512
37	SER	O	2.565	50.303	-16.151
37	SER	OG	-0.008	59.133	-13.079
38	SER	CA	6.261	59.985	-14.607
38	SER	O	6.963	59.251	-15.205
38	SER	OG	5.376	59.065	-12.234
39	HIS	CA	6.637	56.576	-15.291
39	HIS	O	5.730	55.070	-17.619
39	HIS	CG	0.016	56.609	-16.656
39	HIS	CD2	0.769	56.365	-13.309
39	HIS	OE2	9.906	53.910	-13.000
40	PRO	CA	7.900	56.097	-18.031
40	PRO	O	0.032	55.097	-20.370
40	PRO	CG	10.853	57.605	-17.902
41	ASP	O	0.601	56.320	-18.605
41	ASP	OD1	10.325	51.395	-20.629
41	ASP	CO	9.790	52.239	-10.226
41	ASP	C	7.311	52.163	-18.039
42	LEU	O	6.105	52.003	-10.550
42	LEU	C	3.926	52.987	-19.376
42	LEU	CO	6.621	52.150	-17.000
42	LEU	CD1	6.535	51.546	-16.501
43	LYS	O	3.010	52.135	-19.066
43	LYS	C	0.637	52.196	-20.010
43	LYS	CO	2.021	52.309	-22.169
43	LYS	CD	0.990	52.062	-24.339
43	LYS	OZ	0.337	51.757	-20.610
44	VAL	CA	-1.687	52.039	-20.765
44	VAL	O	-2.629	53.906	-20.634
44	VAL	CG1	-2.726	52.941	-16.502
45	ALA	O	-3.696	51.951	-19.071
45	ALA	C	-5.061	52.567	-20.053
45	ALA	CO	-6.031	50.900	-21.309
46	GLY	CA	-7.002	52.037	-10.001
46	GLY	O	-5.930	52.006	-16.035
47	GLY	CA	-0.816	52.266	-16.300
47	GLY	O	-9.000	53.601	-16.105
40	ALA	CA	-10.255	52.070	-11.302
40	ALA	O	-9.066	53.720	-0.725
49	SER	O	-10.169	53.967	-9.037
49	SER	C	-10.967	52.906	-6.703
49	SER	CO	-9.892	56.500	-7.829
50	MET	O	-10.035	52.007	-5.932
50	MET	C	-11.663	51.962	-3.961
50	MET	CO	-12.012	50.010	-6.996
50	MET	SD	-13.600	69.009	-7.256
51	VAL	O	-10.627	52.760	-3.622
51	VAL	C	-10.638	56.562	-2.907
51	VAL	CO	-0.663	53.355	-2.000
51	VAL	CG2	-7.766	53.015	-2.082
52	PRO	CA	-12.372	55.933	-0.021
52	PRO	O	-11.771	50.220	-0.029
52	PRO	CG	-13.503	56.103	0.005
53	SER	O	-10.642	50.006	0.290
53	SER	C	-0.620	50.268	-0.326
53	SER	CO	-9.004	57.707	2.060
54	GLU	O	-0.256	57.523	-1.393
54	GLU	C	-9.767	57.303	-3.705
54	GLU	CO	-6.136	56.590	-2.356
54	GLU	CA	-6.066	56.067	-0.090
36	ASP	CO	3.712	55.720	-10.516
36	ASP	OD1	3.755	57.976	-11.629
37	SER	O	1.306	56.022	-13.311
37	SER	C	2.377	50.095	-16.969
37	SER	CO	-0.093	50.069	-16.700
38	SER	O	3.163	50.616	-16.001
38	SER	C	3.666	50.705	-16.992
38	SER	CO	6.762	60.635	-13.390
39	HIS	O	5.656	57.300	-16.092
39	HIS	C	6.601	56.681	-16.770
39	HIS	CO	6.637	55.283	-16.515
39	HIS	OD1	0.795	56.356	-15.561
39	HIS	CE1	9.978	53.930	-15.130
40	PRO	O	7.007	56.036	-17.307
40	PRO	C	0.156	55.200	-19.357
40	PRO	CD	9.267	57.533	-19.361
40	PRO	CD	0.900	57.652	-16.776
41	ASP	OD2	11.160	50.399	-10.660
41	ASP	CG	10.673	51.307	-19.211
41	ASP	CA	0.665	52.959	-10.966
41	ASP	O	7.396	50.967	-10.977
42	LEU	CA	6.092	52.167	-10.666
42	LEU	O	3.993	56.163	-19.690
42	LEU	CG	5.102	51.363	-15.966
42	LEU	CD2	5.273	60.077	-16.350
43	LYS	CA	2.093	52.605	-20.721
43	LYS	O	0.504	50.920	-19.020
43	LYS	CG	0.605	52.636	-22.910
43	LYS	CE	-0.308	52.506	-25.260
44	VAL	O	-0.191	53.035	-19.690
44	VAL	C	-2.571	52.007	-19.731
44	VAL	CO	-2.608	53.351	-17.303
44	VAL	CG2	-0.197	53.196	-16.553
45	ALA	CA	-6.619	51.077	-20.010
45	ALA	O	-6.703	53.005	-20.703
46	GLY	O	-5.910	52.356	-10.760
46	GLY	C	-6.907	52.663	-16.530
47	GLY	O	-0.002	52.650	-15.703
47	GLY	C	-9.279	52.757	-13.572
40	ALA	O	-9.221	52.666	-12.330
40	ALA	C	-9.790	52.675	-9.960
40	ALA	CO	-11.550	52.100	-11.617
49	SER	CA	-9.792	53.355	-7.652
49	SER	O	-11.072	53.677	-6.980
49	SER	OG	-0.077	56.255	-5.650
50	MET	CA	-11.052	51.969	-6.976
50	MET	O	-11.997	51.390	-2.575
50	MET	CG	-11.912	69.663	-6.309
50	MET	CE	-12.080	50.111	-0.983
51	VAL	CA	-9.060	53.170	-2.067
51	VAL	O	-10.237	55.637	-2.602
51	VAL	CG1	-7.002	53.579	-0.631
52	PRO	O	-11.621	56.693	-1.056
52	PRO	C	-11.690	57.123	-0.660
52	PRO	CO	-13.600	55.906	0.266
52	PRO	CD	-12.366	53.620	-0.175
53	SER	CA	-9.300	57.902	0.602
53	SER	O	-7.690	59.226	-0.030
53	SER	OG	-0.256	56.521	2.127
54	GLU	CA	-7.206	57.660	-2.621
54	GLU	O	-7.533	56.263	-6.370
54	GLU	CG	-5.200	56.090	-0.927
54	GLN	OD1	-5.066	56.606	-1.060

54	GLU	OE2	-3.900	55.777	0.271
55	THR	CA	-9.433	58.121	-5.441
55	THR	O	-9.633	57.919	-7.010
55	THR	OG1	-9.005	60.510	-5.410
56	ASN	N	-7.602	58.603	-6.077
56	ASN	OD1	-5.075	58.967	-10.337
56	ASN	CG	-5.090	59.694	-8.200
56	ASN	C	-6.012	57.094	-8.305
57	PRO	N	-6.362	56.261	-9.250
57	PRO	CD	-7.304	56.633	-10.272
57	PRO	CA	-5.679	54.961	-9.332
57	PRO	O	-3.509	54.120	-9.945
58	PHE	CA	-2.747	56.577	-11.222
58	PHE	O	-0.635	57.497	-10.600
58	PHE	CG	-3.903	56.960	-13.357
58	PHE	CD2	-5.211	57.630	-13.659
58	PHE	CE2	-6.194	57.095	-14.276
59	GLN	N	-2.064	57.119	-8.990
59	GLN	C	-0.087	56.603	-7.000
59	GLN	CB	-1.062	58.640	-7.009
59	GLN	CD	-1.790	60.157	-5.150
59	GLN	NE2	-2.959	59.605	-6.742
60	ASP	CA	0.051	54.792	-6.304
60	ASP	O	2.027	55.550	-5.231
60	ASP	CG	2.077	52.530	-6.300
60	ASP	OD2	2.915	51.041	-7.030
61	ASN	ND2	-1.364	57.747	-2.347
61	ASN	CG	-0.040	57.670	-2.399
61	ASN	CA	1.557	55.734	-2.700
61	ASN	O	2.933	54.062	-0.902
62	ASN	CA	2.077	52.340	-1.709
62	ASN	O	4.951	51.313	-1.770
62	ASN	CG	2.371	50.103	-0.697
62	ASN	ND2	2.622	50.200	0.601
63	SER	CA	5.109	51.696	-4.709
63	SER	O	5.573	49.798	-6.269
63	SER	OG	6.071	50.690	-3.610
64	HIS	CA	3.994	48.055	-4.935
64	HIS	O	3.061	46.974	-7.108
64	HIS	CG	3.144	46.021	-3.726
64	HIS	CD2	4.856	45.194	-3.135
64	HIS	NE2	3.556	43.920	-3.360
65	GLY	CA	1.552	48.264	-7.030
65	GLY	O	2.230	48.870	-10.136
66	THR	CA	4.064	50.117	-9.954
66	THR	O	5.333	48.709	-11.461
66	THR	OG1	3.637	52.425	-9.406
67	HIS	N	5.605	48.443	-9.276
67	HIS	C	6.091	46.141	-10.143
67	HIS	CB	7.300	47.071	-8.064
67	HIS	ND1	8.598	46.907	-8.276
67	HIS	CE1	9.057	46.691	-8.299
68	VAL	N	4.092	45.749	-9.731
68	VAL	C	3.856	44.860	-11.748
68	VAL	CB	2.939	44.252	-9.306
68	VAL	CG2	3.319	43.785	-8.000
69	ALA	CA	3.037	46.460	-13.629
69	ALA	O	4.020	45.913	-15.965
70	GLY	N	5.340	46.702	-13.914
70	GLY	C	7.040	45.370	-15.021
71	THR	N	6.020	44.431	-14.130
71	THR	C	6.224	42.506	-15.543
71	THR	CB	7.119	42.078	-13.191
55	THR	N	-0.971	58.291	-4.249
55	THR	C	-0.764	58.130	-6.779
55	THR	CB	-10.506	59.208	-5.303
55	THR	CG2	-11.432	59.143	-4.817
56	ASN	ND2	-6.930	61.179	-9.001
56	ASN	CB	-5.273	59.925	-9.555
56	ASN	CA	-6.762	58.425	-8.200
56	ASN	O	-5.184	56.066	-7.470
57	PRO	CG	-7.123	55.257	-11.177
57	PRO	CB	-6.644	54.178	-10.235
57	PRO	C	-4.381	55.002	-9.966
58	PHE	N	-3.990	56.262	-10.491
58	PHE	C	-1.712	57.129	-10.253
58	PHE	CB	-2.943	57.582	-12.423
58	PHE	CD1	-3.756	55.708	-14.059
58	PHE	CE1	-4.722	55.255	-14.920
58	PHE	CZ	-5.949	55.939	-15.051
59	GLN	CA	-1.172	57.503	-7.934
59	GLN	O	-1.639	56.083	-6.115
59	GLN	CG	-0.942	59.261	-6.034
59	GLN	OE1	-1.604	61.200	-4.836
60	ASP	N	0.418	55.095	-7.211
60	ASP	C	1.631	55.267	-5.098
60	ASP	CB	1.596	53.744	-7.100
60	ASP	OD1	1.766	52.337	-5.190
61	ASN	N	0.959	55.265	-3.950
61	ASN	OD1	0.666	58.566	-2.075
61	ASN	CB	0.531	56.401	-1.704
61	ASN	C	2.291	54.632	-1.948
62	ASN	N	2.218	53.434	-2.460
62	ASN	C	4.124	51.093	-2.479
62	ASN	CB	1.703	51.319	-1.421
62	ASN	OD1	2.633	49.077	-1.363
63	SER	N	4.352	52.184	-3.761
63	SER	C	5.071	50.256	-5.289
63	SER	CB	6.523	51.950	-6.012
64	HIS	N	4.282	49.475	-4.639
64	HIS	C	3.366	47.759	-6.261
64	HIS	CB	3.104	47.581	-3.747
64	HIS	ND1	2.107	45.247	-6.241
64	HIS	CE1	2.416	43.966	-4.054
65	GLY	N	2.207	48.620	-6.587
65	GLY	C	2.392	48.636	-9.037
66	THR	N	3.233	49.659	-8.832
66	THR	C	5.009	49.009	-10.291
66	THR	CB	4.744	51.511	-9.667
66	THR	CG2	5.536	52.070	-10.849
67	HIS	CA	6.703	47.361	-9.458
67	HIS	O	6.649	45.630	-11.150
67	HIS	CG	8.595	46.275	-8.140
67	HIS	CD2	9.904	46.670	-8.876
67	HIS	NE2	10.678	45.514	-8.106
68	VAL	CA	4.242	46.607	-10.206
68	VAL	O	4.314	43.942	-12.535
68	VAL	CG1	1.960	43.260	-10.820
69	ALA	N	3.373	46.049	-12.113
69	ALA	C	4.393	46.398	-14.411
69	ALA	CB	2.332	47.051	-13.306
70	GLY	CA	6.595	46.085	-14.678
70	GLY	O	7.604	45.154	-16.119
71	THR	CA	7.177	43.019	-16.466
71	THR	O	6.602	41.020	-16.095
71	THR	OG1	8.391	42.592	-12.398

71	THR	CG2	7.276	60.903	-19.596	72	VAL	O	6.090	62.007	-19.627
72	VAL	CA	3.976	62.491	-16.604	72	VAL	C	6.312	63.004	-17.031
72	VAL	O	6.341	62.300	-10.060	72	VAL	CD	2.916	62.067	-16.005
72	VAL	CG1	1.512	62.600	-17.170	72	VAL	CG2	2.142	62.327	-14.723
73	ALA	O	6.936	66.417	-17.080	73	ALA	CA	6.907	63.091	-19.167
73	ALA	C	5.633	66.333	-19.355	73	ALA	O	5.062	67.100	-20.216
73	ALA	CD	3.107	63.441	-19.433	74	ALA	O	6.546	66.629	-18.635
74	ALA	CA	7.670	67.591	-18.059	74	ALA	C	7.740	67.040	-20.342
74	ALA	O	7.959	66.649	-21.054	74	ALA	CO	8.653	67.446	-17.925
75	LEU	O	7.658	68.704	-21.039	75	LEU	CA	7.012	68.960	-22.456
75	LEU	C	9.192	68.560	-22.966	75	LEU	O	10.162	68.790	-22.253
75	LEU	CO	7.540	70.471	-22.089	75	LEU	CG	6.123	70.913	-22.370
75	LEU	CD1	6.879	72.436	-22.308	75	LEU	CD2	5.096	70.662	-23.405
76	ASN	O	9.147	68.103	-24.169	76	ASN	OD2	12.305	66.632	-26.304
76	ASN	OD1	10.950	65.048	-27.920	76	ASN	CG	11.195	66.276	-26.082
76	ASN	CO	10.010	66.651	-25.900	76	ASN	CA	10.359	67.730	-24.930
76	ASN	C	10.783	69.040	-25.643	76	ASN	O	10.157	69.679	-26.619
77	ASN	O	11.004	69.664	-25.071	77	ASN	CA	12.228	70.957	-25.601
77	ASN	C	13.787	71.029	-25.340	77	ASN	O	14.364	69.979	-25.333
77	ASN	CO	11.335	72.876	-25.117	77	ASN	CG	11.250	72.027	-23.616
77	ASN	OD1	12.032	73.346	-22.917	77	ASN	OD2	10.294	72.741	-23.025
78	SER	O	14.125	72.267	-25.164	78	SER	CA	15.513	72.614	-24.986
78	SER	C	15.010	72.742	-23.436	78	SER	O	16.902	73.071	-23.164
78	SER	CO	15.985	73.941	-25.587	78	SER	OG	15.926	73.078	-26.999
79	ILE	O	14.050	72.565	-22.529	79	ILE	CA	15.155	72.704	-21.120
79	ILE	C	14.617	71.603	-20.230	79	ILE	O	13.043	70.061	-20.679
79	ILE	CO	14.471	74.174	-20.697	79	ILE	CG1	12.945	74.032	-20.014
79	ILE	CG2	14.997	75.329	-21.612	79	ILE	CD1	12.135	75.176	-20.155
80	GLY	O	14.995	71.760	-18.901	80	GLY	CA	14.476	70.960	-17.913
80	GLY	C	14.612	69.440	-18.219	80	GLY	O	15.719	68.994	-18.944
81	VAL	O	13.513	68.766	-17.900	81	VAL	CA	13.411	67.206	-18.061
81	VAL	C	12.511	66.919	-19.217	81	VAL	O	12.268	67.739	-20.117
81	VAL	CO	13.801	66.755	-16.677	81	VAL	CG1	14.038	67.004	-15.573
81	VAL	CG2	11.630	67.281	-16.231	82	LEU	O	12.126	65.645	-19.216
82	LEU	CA	11.312	65.820	-20.256	82	LEU	C	10.998	64.020	-19.518
82	LEU	O	10.058	63.356	-18.680	82	LEU	CO	12.206	64.219	-21.229
82	LEU	CG	11.430	63.560	-22.366	82	LEU	CD1	10.796	64.657	-23.223
82	LEU	CD2	12.359	62.675	-23.192	83	GLY	O	9.131	64.100	-19.016
83	GLY	CA	8.133	63.921	-19.134	83	GLY	C	8.027	62.011	-19.925
83	GLY	O	8.946	61.022	-21.026	84	VAL	O	7.272	61.112	-19.203
84	VAL	CA	6.973	59.007	-19.080	84	VAL	C	6.164	60.038	-21.140
84	VAL	O	6.424	59.472	-22.194	84	VAL	CO	6.256	58.920	-18.041
84	VAL	CG1	5.608	57.677	-19.557	84	VAL	CG2	7.190	58.587	-17.765
85	ALA	O	5.156	60.926	-21.024	85	ALA	CA	6.217	61.194	-22.150
85	ALA	C	4.213	62.603	-22.396	85	ALA	O	3.268	63.601	-22.030
85	ALA	CO	2.046	60.663	-21.748	86	PRO	N	5.248	63.106	-23.059
86	PRO	CA	5.413	64.635	-23.205	86	PRO	C	6.321	65.371	-23.947
86	PRO	O	6.291	66.695	-23.049	86	PRO	CO	6.022	64.704	-23.013
86	PRO	CG	7.038	63.466	-24.546	86	PRO	CB	6.977	62.648	-23.636
87	SER	O	3.540	66.676	-24.769	87	SER	CA	2.609	65.324	-25.529
87	SER	C	3.103	65.132	-24.097	87	SER	O	0.162	65.913	-25.619
87	SER	CO	2.601	66.977	-26.927	87	SER	OG	3.591	65.163	-27.503
88	ALA	O	1.817	66.964	-23.742	88	ALA	CO	-0.163	63.910	-21.020
88	ALA	CA	-0.273	66.953	-23.884	88	ALA	C	-0.090	65.717	-22.698
88	ALA	O	-0.174	66.717	-22.435	89	SER	N	-2.219	69.691	-22.670
89	SER	OG	-4.166	67.182	-24.208	89	SER	CB	-4.340	68.903	-22.090
89	SER	CA	-3.001	66.067	-22.227	89	SER	C	-3.106	66.708	-20.727
89	SER	O	-3.793	65.064	-20.209	90	LEU	O	-2.446	67.656	-20.037
90	LEU	CA	-2.370	67.667	-10.593	90	LEU	C	-3.603	60.630	-17.064
90	LEU	O	-3.902	69.606	-10.215	90	LEU	CG	-0.931	60.273	-10.476
90	LEU	CG	-0.233	67.051	-17.174	90	LEU	CD1	-0.026	66.361	-17.219
90	LEU	CD2	1.168	60.524	-17.847	91	TYR	O	-4.266	67.944	-16.930
91	TYR	CA	-5.250	60.070	-16.137	91	TYR	C	-4.073	60.738	-14.605

91	TYR	O	-6.696	67.749	-14.023	91	TYR	CD	-6.606	68.093	-14.314
91	TYR	CG	-7.094	68.237	-17.741	91	TYR	CD1	-6.995	67.415	-10.755
91	TYR	CD2	-7.971	69.275	-18.149	91	TYR	CE1	-6.905	67.972	-28.098
91	TYR	CE2	-8.315	68.421	-19.692	91	TYR	CZ	-7.794	68.902	-20.463
91	TYR	OH	-8.102	68.752	-21.764	92	ALA	O	-4.095	69.958	-14.104
92	ALA	CA	-6.949	68.199	-12.707	92	ALA	C	-5.023	68.033	-11.903
92	ALA	B	-6.723	68.898	-12.858	92	ALA	CB	-3.997	68.621	-12.400
93	VAL	O	-5.959	68.993	-11.229	93	VAL	CA	-7.103	68.054	-10.325
93	VAL	C	-8.788	69.014	-8.099	93	VAL	B	-6.181	67.993	-8.372
93	VAL	CB	-7.957	67.935	-10.631	93	VAL	CG1	-8.213	67.408	-9.725
93	VAL	CG2	-8.195	67.370	-12.072	94	LYS	O	-6.907	68.217	-8.327
94	LYS	CA	-6.378	68.464	-6.999	94	LYS	C	-7.331	69.905	-5.094
94	LYS	B	-8.458	68.408	-5.783	94	LYS	CB	-6.051	68.976	-6.818
94	LYS	CG	-5.394	68.328	-5.667	94	LYS	CD	-6.068	68.785	-5.502
94	LYS	CE	-4.399	68.288	-4.199	94	LYS	OZ	-3.735	68.544	-4.387
95	VAL	O	-6.989	69.871	-3.826	95	VAL	CA	-7.646	68.457	-3.928
95	VAL	C	-6.919	68.499	-2.568	95	VAL	B	-7.425	68.156	-1.581
95	VAL	CB	-8.184	67.038	-4.319	95	VAL	CG1	-8.068	66.852	-5.619
95	VAL	CG2	-6.988	66.108	-4.832	96	LEU	O	-5.676	68.974	-2.604
96	LEU	CA	-6.782	69.103	-1.406	96	LEU	C	-4.331	68.559	-1.321
96	LEU	B	-3.962	81.121	-2.336	96	LEU	CB	-3.909	68.241	-1.573
96	LEU	CG	-3.593	66.799	-2.872	96	LEU	CD1	-2.287	66.184	-2.163
96	LEU	CD2	-4.489	66.882	-1.045	97	GLY	O	-4.326	68.975	-0.086
97	GLY	CA	-3.098	82.387	0.287	97	GLY	C	-2.363	82.437	0.305
97	GLY	O	-1.619	81.463	0.165	98	ALA	O	-1.954	83.648	0.758
98	ALA	CB	-0.428	85.678	1.510	98	ALA	CA	-0.563	84.068	0.965
98	ALA	C	0.188	83.118	1.917	98	ALA	B	1.393	82.921	1.663
99	ASP	O	-0.584	82.573	2.912	99	ASP	OD2	-2.631	81.042	6.151
99	ASP	OD1	-2.738	80.982	6.003	99	ASP	CG	-2.003	81.131	5.060
99	ASP	CB	-0.648	81.603	5.175	99	ASP	CA	0.101	81.618	3.055
99	ASP	C	0.166	80.165	3.328	99	ASP	B	0.735	69.313	4.029
100	GLY	O	-0.424	69.083	2.168	100	GLY	CA	-8.343	68.521	3.615
100	GLY	C	-1.520	67.651	2.882	100	GLY	B	-1.649	66.522	3.479
101	SER	O	-2.342	68.128	2.988	101	SER	CA	-3.542	67.388	3.315
101	SER	C	-4.759	67.094	2.532	101	SER	B	-4.758	68.972	1.907
101	SER	CB	-3.716	67.447	4.017	101	SER	OG	-4.411	68.634	5.209
102	GLY	O	-5.021	67.092	2.577	102	GLY	CA	-7.077	67.422	1.096
102	GLY	C	-8.166	66.536	2.520	102	GLY	B	-7.088	65.431	3.038
103	GLN	O	-9.377	67.058	2.698	103	GLN	CA	-10.535	66.297	3.029
103	GLN	C	-10.963	65.232	2.022	103	GLN		-10.779	69.402	8.017
103	GLN	CB	-11.671	67.387	3.274	103	GLN	CG	-11.368	68.885	6.506
103	GLN	CD	-12.360	69.186	6.915	103	GLN	OE1	-12.159	69.816	5.982
103	GLN	OE2	-13.619	69.197	6.112	104	TYR	O	-11.611	64.141	2.451
104	TYR	CA	-12.068	63.326	1.508	104	TYR	C	-13.031	63.698	0.473
104	TYR	B	-12.939	63.276	-0.607	104	TYR	CG	-12.697	61.866	2.143
104	TYR	CG	-11.629	68.029	2.472	104	TYR	CD1	-11.019	59.709	3.377
104	TYR	CD2	-10.379	60.959	1.068	104	TYR	CE1	-10.089	58.005	3.787
104	TYR	CE2	-9.352	60.857	2.171	104	TYR	CZ	-9.564	59.022	3.001
104	TYR	OH	-8.401	80.191	3.326	105	SER	O	-13.989	64.572	0.903
105	SER	CA	-14.077	65.166	-0.034	105	SER	C	-14.172	65.978	-1.159
105	SER	O	-14.759	65.935	-2.250	105	SER	CB	-15.008	66.121	0.601
105	SER	OG	-15.289	67.039	1.458	106	TRP	O	-13.079	66.625	-0.034
106	TRP	CA	-12.421	67.391	-1.940	106	TRP	C	-11.095	66.436	-3.012
106	TRP	B	-12.021	66.640	-4.245	106	TRP	CG	-11.321	68.256	-1.355
106	TRP	CG	-11.645	69.111	-0.286	106	TRP	CD1	-12.062	69.926	0.264
106	TRP	CD2	-10.650	69.012	0.901	106	TRP	NE1	-12.691	70.350	1.369
106	TRP	CE2	-11.359	70.573	1.561	106	TRP	CE3	-9.275	69.052	0.576
106	TRP	CZ2	-10.671	71.310	2.580	106	TRP	CZ3	-8.960	70.963	1.525
106	TRP	CH2	-9.293	71.291	2.455	107	ILE	O	-11.390	68.330	-2.401
107	ILE	CA	-10.765	64.258	-3.325	107	ILE	C	-11.095	68.594	-6.198
107	ILE	B	-11.695	63.674	-5.390	107	ILE	CG	-9.066	68.103	-2.923
107	ILE	CG1	-8.634	63.784	-1.936	107	ILE	CG2	-9.632	61.938	-3.301
107	ILE	CD1	-8.253	62.998	-0.627	108	ILE	O	-12.994	63.292	-3.577

100	ILE	CA	-14.116	62.722	-4.321	108	ILE	C	-14.630	63.694	-5.386
108	ILE	O	-14.094	63.320	-6.552	108	ILE	CO	-15.246	62.265	-3.320
100	ILE	CG1	-14.726	61.077	-2.482	109	ILE	CG2	-16.960	62.024	-4.095
100	ILE	CD1	-15.432	60.045	-1.131	109	ASN	N	-14.751	66.990	-4.901
109	ASN	CA	-15.284	66.010	-5.916	109	ASN	C	-14.232	66.067	-7.004
109	ASN	O	-14.668	66.272	-8.235	109	ASN	CO	-15.208	67.359	-5.207
109	ASN	CG	-16.520	67.406	-6.353	109	ASN	OD1	-17.695	66.695	-6.646
109	ASN	OD2	-16.633	68.447	-3.442	110	GLY	N	-12.981	68.900	-6.774
110	GLY	CA	-11.952	69.917	-7.065	110	GLY	C	-12.100	64.712	-8.012
110	GLY	O	-11.929	64.929	-10.036	111	ILE	N	-12.379	63.539	-8.246
111	ILE	CA	-12.603	62.334	-9.099	111	ILE	C	-13.059	62.368	-9.942
111	ILE	O	-13.921	62.304	-11.140	111	ILE	CO	-12.736	60.948	-8.364
111	ILE	CG1	-11.621	60.501	-7.655	111	ILE	CG2	-13.122	59.791	-9.347
111	ILE	CD1	-11.500	59.706	-6.336	112	GLU	N	-14.093	63.075	-9.208
112	GLU	CA	-16.138	63.376	-10.046	112	GLU	C	-15.072	64.347	-11.171
112	GLU	O	-16.667	64.136	-12.266	112	GLU	CO	-17.229	63.099	-9.141
112	GLU	CG	-17.047	62.917	-8.135	112	GLU	CD	-18.726	61.026	-8.605
112	GLU	OE1	-19.041	60.066	-8.816	112	GLU	OE2	-19.123	61.920	-9.066
113	TRP	N	-15.896	65.483	-10.971	113	TRP	CA	-16.756	66.600	-12.000
113	TRP	C	-14.076	65.663	-13.148	113	TRP	O	-14.319	68.932	-14.332
113	TRP	CO	-13.882	67.553	-11.434	113	TRP	CG	-13.406	60.556	-12.601
113	TRP	CD1	-14.140	69.736	-12.601	113	TRP	CD2	-12.641	60.952	-13.463
113	TRP	NE1	-13.597	70.443	-13.723	113	TRP	CE2	-12.545	69.761	-14.215
113	TRP	CE3	-11.451	67.645	-13.009	113	TRP	CZ2	-11.696	70.045	-15.274
113	TRP	CZ3	-10.610	67.099	-14.079	113	TRP	CH2	-10.752	69.074	-15.683
114	ALA	N	-13.889	64.001	-12.032	114	ALA	CA	-12.333	64.065	-13.074
114	ALA	C	-13.199	63.179	-14.752	114	ALA	O	-12.963	63.074	-15.970
114	ALA	CO	-11.299	63.192	-13.140	115	ILE	N	-14.174	62.360	-14.119
115	ILE	CA	-15.070	61.640	-14.097	115	ILE	C	-15.920	62.605	-15.056
115	ILE	O	-16.077	62.225	-17.070	115	ILE	CO	-16.000	60.048	-13.922
115	ILE	CG1	-15.210	59.036	-13.843	115	ILE	CG2	-17.151	60.160	-14.795
115	ILE	CD1	-16.804	59.411	-11.763	116	ALA	N	-16.934	63.527	-15.267
116	ALA	CA	-17.398	64.448	-16.050	116	ALA	C	-16.706	65.069	-17.270
116	ALA	O	-17.323	65.255	-18.343	116	ALA	CO	-18.011	65.910	-15.151
117	ASN	N	-15.423	65.390	-17.122	117	ASN	CA	-16.953	69.967	-18.139
117	ASN	C	-13.027	64.974	-19.034	117	ASN	O	-12.997	65.636	-19.020
117	ASN	CO	-13.615	66.958	-17.426	117	ASN	CG	-16.688	60.177	-16.939
117	ASN	OD1	-14.565	69.082	-17.773	117	ASN	ND2	-16.931	60.269	-15.736
118	ASN	N	-14.223	63.725	-18.967	118	ASN	CA	-13.760	62.662	-19.032
118	ASN	C	-12.248	62.644	-19.043	118	ASN	O	-11.617	62.389	-20.932
118	ASN	CO	-14.247	62.063	-21.279	118	ASN	CG	-15.737	63.060	-21.395
118	ASN	OD1	-16.518	62.321	-20.759	118	ASN	ND2	-16.136	64.896	-22.133
119	MET	N	-11.606	62.588	-18.675	119	MET	CA	-10.232	62.222	-18.670
119	MET	C	-10.825	60.734	-18.920	119	MET	O	-10.000	39.030	-18.759
119	MET	CO	-9.018	62.461	-17.855	119	MET	CG	-9.000	63.083	-16.902
119	MET	SD	-8.700	64.943	-17.526	119	MET	CE	-9.902	66.061	-18.263
120	ASP	N	-8.984	60.637	-19.504	120	ASP	CA	-8.400	39.110	-20.038
120	ASP	C	-7.022	38.398	-18.056	120	ASP	O	-8.030	37.109	-18.698
120	ASP	CO	-7.555	39.156	-21.236	120	ASP	CG	-8.237	39.730	-22.454
120	ASP	OD1	-7.001	60.706	-23.004	120	ASP	OD2	-9.327	39.135	-22.739
121	VAL	N	-7.871	39.117	-18.115	121	VAL	CA	-6.226	38.601	-16.974
121	VAL	C	-6.296	39.534	-15.706	121	VAL	O	-6.204	68.700	-15.909
121	VAL	CO	-4.755	38.507	-17.496	121	VAL	CG1	-3.758	38.176	-16.427
121	VAL	CG2	-4.707	37.916	-38.046	122	ILE	N	-6.310	38.970	-14.998
122	ILE	CA	-6.240	39.799	-13.397	122	ILE	C	-5.020	39.262	-12.627
122	ILE	O	-6.029	38.012	-12.469	122	ILE	CO	-7.476	39.684	-12.666
122	ILE	CG1	-0.606	60.392	-13.063	122	ILE	CG2	-7.221	39.003	-10.954
122	ILE	CD1	-0.976	39.708	-12.393	123	ASN	N	-6.263	60.222	-12.110
123	ASN	CA	-3.165	39.054	-11.232	123	ASN	C	-3.902	60.684	-9.061
123	ASN	O	-3.708	61.631	-9.033	123	ASN	CG	-1.020	60.470	-11.697
123	ASN	CG	-0.692	60.040	-10.777	123	ASN	OD1	-0.063	39.998	-11.010
123	ASN	ND2	-0.366	58.747	-9.720	124	MET	N	-3.650	39.684	-8.032
124	MET	CA	-3.650	39.973	-7.630	124	MET	C	-2.623	39.603	-6.614

124	MET	O	-2.306	30.300	-6.093	124	MET	CB	-6.943	30.307	-6.090
124	MET	CG	-6.190	40.002	-7.673	124	MET	SD	-7.308	30.473	-6.430
124	MET	CE	-7.040	30.091	-7.042	125	SER	N	-1.484	40.496	-6.083
125	SER	CA	-0.193	40.207	-3.709	125	SER	C	-3.423	40.712	-6.326
125	SER	O	0.233	41.017	-3.003	125	SER	CB	1.023	41.027	-6.323
125	SER	OG	2.444	40.496	-7.373	126	LEU	N	-1.433	40.078	-3.773
126	LEU	CA	-1.042	40.347	-2.306	126	LEU	C	-2.438	30.096	-1.007
126	LEU	O	-2.044	30.136	-2.829	126	LEU	CB	-2.701	41.860	-2.410
126	LEU	CG	-3.000	41.447	-3.333	126	LEU	CD1	-8.278	41.131	-2.370
126	LEU	CD2	-4.170	42.700	-4.073	127	GLY	N	-2.822	30.082	-3.481
127	GLY	CA	-3.033	37.071	0.193	127	GLY	C	-3.176	38.100	2.082
127	GLY	O	-2.646	30.836	2.226	128	GLY	N	-4.121	37.443	2.222
128	GLY	CA	-6.473	37.400	3.042	128	GLY	C	-4.644	36.030	4.104
128	GLY	O	-6.083	38.180	3.276	129	PRO	N	-4.919	38.007	8.482
129	PRO	CA	-6.671	34.823	3.990	129	PRO	C	-8.116	34.006	6.002
129	PRO	O	-6.338	32.807	6.303	129	PRO	CB	-6.063	34.604	7.304
129	PRO	CG	-4.619	36.318	7.727	129	PRO	CD	-6.239	38.070	6.410
130	SER	N	-7.831	33.813	8.932	130	SER	CA	-8.470	34.631	6.823
130	SER	C	-9.210	34.004	6.726	130	SER	O	-8.940	33.001	4.029
130	SER	CB	-9.069	35.351	7.216	130	SER	OG	-8.723	36.626	0.483
131	GLY	N	-10.003	33.967	6.849	131	GLY	CA	-10.024	34.229	3.076
131	GLY	C	-12.265	34.713	3.842	131	GLY	O	-12.493	34.722	6.781
132	SER	N	-13.040	33.830	2.594	132	SER	CA	-14.607	38.433	3.011
132	SER	C	-15.209	34.003	1.936	132	SER	O	-14.799	34.886	0.026
132	SER	CB	-14.899	36.027	3.145	132	SER	OG	-14.093	37.339	1.073
133	ALA	N	-16.867	34.900	2.206	133	ALA	CA	-17.397	34.087	1.326
133	ALA	C	-17.630	34.965	0.897	133	ALA	O	-17.763	34.437	-1.016
133	ALA	CB	-18.866	33.020	1.996	134	ALA	N	-17.603	36.200	0.294
134	ALA	CA	-17.072	37.239	-0.792	134	ALA	C	-18.633	37.369	-1.676
134	ALA	O	-16.701	37.803	-2.069	134	ALA	CB	-18.263	38.606	-0.307
135	LEU	N	-13.470	37.229	-3.046	135	LEU	CA	-14.107	37.266	-1.004
135	LEU	C	-14.130	36.003	-2.705	135	LEU	O	-13.794	36.020	-3.090
135	LEU	CB	-13.830	37.320	-0.790	135	LEU	CG	-11.693	37.130	-1.808
135	LEU	CD1	-11.460	30.415	-2.292	135	LEU	CD2	-10.802	34.087	-0.819
136	LYS	N	-14.899	34.023	-2.173	136	LYS	CA	-14.943	33.997	-3.013
136	LYS	C	-13.844	33.739	-6.190	136	LYS	C	-13.270	33.431	-3.303
136	LYS	CB	-16.903	32.341	-2.106	136	LYS	CG	-14.743	31.067	-3.043
136	LYS	CD	-15.803	20.892	-2.134	136	LYS	CE	-13.743	28.787	-2.778
136	LYS	NZ	-15.308	20.611	-4.160	137	ALA	N	-16.746	34.260	-3.047
137	ALA	CA	-17.793	34.416	-4.003	137	ALA	C	-17.330	38.383	-6.043
137	ALA	O	-17.703	33.840	-7.298	137	ALA	CB	-18.804	34.961	-6.263
138	ALA	N	-16.829	36.301	-8.729	138	ALA	CA	-16.801	37.311	-6.089
138	ALA	C	-14.903	36.096	-7.837	138	ALA	O	-14.905	36.063	-8.762
138	ALA	CB	-13.822	38.967	-3.934	139	VAL	N	-13.980	33.989	-7.027
139	VAL	CA	-12.946	33.291	-7.037	139	VAL	C	-13.628	34.820	-8.728
139	VAL	O	-13.280	34.870	-9.077	139	VAL	CB	-11.030	34.071	-6.960
139	VAL	CG1	-10.919	33.036	-7.066	139	VAL	CG2	-11.898	33.700	-6.213
140	ASP	N	-16.893	33.836	-8.122	140	ASP	CA	-18.274	32.696	-8.929
140	ASP	C	-10.023	33.131	-10.804	140	ASP	O	-16.900	32.870	-11.390
140	ASP	CB	-10.149	31.849	-8.308	140	ASP	CG	-15.300	38.040	-7.306
140	ASP	OD1	-14.170	30.483	-7.202	140	ASP	OD2	-16.130	30.132	-6.329
141	LYS	N	-20.030	34.203	-9.020	141	LYS	CA	-17.373	35.886	-10.060
141	LYS	C	-10.373	38.413	-11.946	141	LYS	O	-16.700	33.240	-13.111
141	LYS	CB	-20.039	38.273	-10.323	141	LYS	CG	-10.004	37.886	-11.306
141	LYS	CD	-19.806	38.107	-10.326	141	LYS	CE	-20.872	39.891	-11.230
141	LYS	NZ	-21.130	48.937	-10.273	142	ALA	N	-15.167	38.040	-11.066
142	ALA	CA	-14.173	36.192	-12.614	142	ALA	C	-13.010	33.818	-13.021
142	ALA	O	-13.778	38.189	-14.785	142	ALA	CB	-12.070	30.097	-11.940
143	VAL	N	-13.802	33.006	-12.032	143	VAL	CA	-13.100	32.783	-13.030
143	VAL	C	-10.346	32.233	-14.606	143	VAL	O	-10.140	31.006	-13.039
143	VAL	CB	-12.831	31.073	-12.716	143	VAL	CG1	-12.308	30.370	-13.401
143	VAL	CG2	-11.383	32.193	-12.814	144	ALA	N	-13.331	32.830	-13.073
144	ALA	CA	-10.966	31.034	-14.041	144	ALA	C	-10.920	32.881	-13.061

366	ALA	C	-17.30C	32.363	-30.939
369	SER	A	-16.307	33.940	-13.706
360	SER	C	-15.609	34.779	-17.029
369	SER	CD	-17.016	36.376	-30.636
366	GLY	N	-16.577	33.006	-17.965
366	GLY	C	-12.373	34.691	-10.305
367	VAL	N	-12.130	35.102	-17.304
347	VAL	C	-9.030	34.036	-16.323
367	VAL	CD	-11.192	36.077	-13.009
367	VAL	CG2	-12.360	37.935	-14.230
360	VAL	CA	-7.602	34.33C	-16.008
360	VAL	B	-6.04C	36.133	-14.700
360	VAL	CG1	-5.079	33.603	-16.201
369	VAL	N	-7.390	36.355	-13.331
369	VAL	C	-5.700	34.305	-11.613
369	VAL	CD	-8.324	34.090	-11.315
369	VAL	CG2	-9.636	35.306	-12.006
350	VAL	CA	-3.393	36.907	-10.901
350	VAL	B	-3.992	36.770	-9.600
350	VAL	CG1	-8.970	34.633	-11.401
351	ALA	N	-2.860	34.940	-0.595
351	ALA	C	-1.000	35.036	-6.657
351	ALA	CO	-3.557	35.390	-6.307
352	ALA	CA	0.714	35.630	-9.132
352	ALA	O	-0.720	34.666	-3.687
353	ALA	N	1.125	33.302	-3.912
353	ALA	C	0.931	32.725	-1.931
353	ALA	CD	1.750	31.030	-3.195
354	GLY	CA	2.063	34.211	0.123
354	GLY	O	4.109	33.267	-0.110
355	ASN	CA	5.344	34.707	2.037
355	ASN	O	6.101	34.029	4.395
355	ASN	CG	5.099	36.702	0.900
355	ASN	ND2	5.634	37.065	0.352
356	GLU	CA	6.633	32.537	4.07C
356	GLU	O	5.376	30.637	0.222
356	GLU	CG	2.691	32.662	0.360
356	GLU	OE1	1.766	34.322	5.312
357	GLY	N	6.309	31.057	4.227
357	GLY	C	6.593	30.622	4.593
358	THR	N	7.147	27.703	5.302
358	THR	OG1	0.707	25.607	6.217
358	THR	CA	6.952	26.607	5.702
358	THR	B	6.670	27.355	7.077
359	SER	OG	3.161	25.906	28.535
359	SER	CA	6.035	25.210	0.095
359	SER	B	3.339	23.201	9.030
360	GLY	CA	3.636	21.306	0.005
36C	GLY	B	6.000	21.326	0.595
361	SER	CA	2.636	19.777	7.956
361	SER	B	0.696	20.367	5.009
361	SER	OG	1.036	10.020	0.505
362	SER	CA	0.167	22.729	7.113
362	SER	B	1.533	23.040	5.506
362	SER	OG	0.206	23.091	9.60C
363	SER	CA	-0.611	26.750	3.990
363	SER	O	-1.070	26.560	3.506
363	SER	OG	-1.902	25.710	2.531
364	THR	CA	0.609	30.360	6.312
364	THR	B	0.605	30.302	5.270
364	THR	OG1	2.906	30.202	5.692
365	VAL	N	-0.515	30.762	2.195
365	VAL	C	-2.020	30.560	3.697

366	ALA	CB	-17.062	31.960	-13.700
365	SER	CA	-16.602	36.017	-16.706
365	SER	O	-15.910	35.311	-10.093
365	SER	OG	-15.552	36.935	-10.069
366	GLY	CA	-13.610	33.709	-10.675
366	GLY	O	-11.620	34.306	-10.366
367	VAL	CA	-10.074	35.036	-16.912
367	VAL	O	-10.571	35.991	-10.606
367	VAL	CG1	-9.096	37.003	-19.070
368	VAL	N	-0.503	35.010	-16.603
368	VAL	C	-7.557	36.907	-16.901
360	VAL	CB	-6.273	36.526	-16.950
360	VAL	CG2	-6.590	33.632	-10.262
369	VAL	CA	-6.007	36.065	-12.369
369	VAL	O	-5.026	35.173	-11.639
369	VAL	CG1	-7.093	35.619	-10.009
350	VAL	N	-6.733	35.561	-11.604
350	VAL	C	-3.357	35.625	-9.099
350	VAL	CB	-2.576	35.505	-11.051
350	VAL	CG2	-2.070	34.063	-13.501
351	ALA	CA	-2.561	35.582	-7.207
351	ALA	O	-0.610	35.009	-6.906
352	ALA	N	-0.690	35.907	-5.022
352	ALA	C	0.386	34.320	-6.100
352	ALA	CB	2.266	36.087	-6.806
353	ALA	CA	0.060	32.358	-2.963
353	ALA	B	0.317	32.392	-0.599
354	GLY	N	1.027	33.693	-1.366
354	GLY	C	3.519	36.069	0.599
355	ASN	N	5.050	36.700	1.560
355	ASN	C	5.399	34.850	3.662
355	ASN	CB	6.888	36.190	1.004
355	ASN	OD1	6.123	36.069	-0.534
356	GLU	N	4.711	33.360	3.675
356	GLU	C	5.322	31.320	0.303
356	GLU	CD	3.209	31.000	0.103
356	GLU	CD	2.594	33.991	6.270
356	GLU	OE2	5.106	36.056	7.166
357	GLY	CA	7.506	30.917	6.507
357	GLY	O	5.616	30.566	6.009
358	THR	CG2	0.870	29.596	5.059
358	THR	CB	7.064	28.566	5.206
358	THR	C	6.390	26.609	7.107
359	SER	N	5.530	20.661	7.697
359	SER	CD	3.673	26.300	0.312
359	SER	C	4.696	23.730	0.566
360	GLY	N	5.076	22.067	0.035
360	GLY	C	6.590	21.040	7.700
361	SER	N	3.035	20.310	0.116
361	SER	C	2.677	20.700	6.706
361	SER	CD	2.566	20.393	7.371
362	SER	N	3.503	23.061	7.609
362	SER	C	0.630	23.002	5.060
362	SER	CD	-0.815	25.600	5.362
363	SER	N	-0.670	25.021	5.307
363	SER	C	-0.661	26.377	6.513
363	SER	CD	-1.090	26.662	5.311
364	THR	N	0.307	26.902	5.002
364	THR	C	0.305	29.206	5.306
364	THR	CD	2.095	30.910	6.030
364	THR	CG2	2.007	27.610	6.081
365	VAL	CA	-0.990	29.962	3.016
365	VAL	O	-2.920	30.102	2.809

165	VAL	CB	-1.339	30.024	-0.381
165	VAL	CG2	-3.216	27.716	-0.695
166	GLY	CA	-2.943	32.770	3.826
166	GLY	O	-6.126	32.306	-0.396
167	TYR	CA	-6.823	34.066	0.133
167	TYR	O	-8.476	36.203	0.004
167	TYR	CG	-7.701	32.064	1.709
167	TYR	CD2	-8.710	32.116	1.393
167	TYR	CE2	-9.068	30.935	3.099
167	TYR	OH	-8.083	29.601	3.630
168	PRO	CG	-6.043	36.376	-3.938
168	PRO	CD	-7.964	35.344	-3.903
168	PRO	C	-6.398	33.336	-3.270
169	GLY	CA	-8.006	33.393	-3.109
169	GLY	C	-6.937	30.702	-3.470
170	LYS	CA	-8.402	30.879	-2.283
170	LYS	C	-7.833	28.773	-2.816
170	LYS	CB	-6.246	29.204	-8.326
170	LYS	CD	-6.230	28.209	2.031
170	LYS	NZ	-4.219	27.463	3.219
171	TYR	CA	-9.812	29.863	-3.839
171	TYR	O	-7.760	30.714	-8.928
171	TYR	CG	-10.497	30.004	-3.047
171	TYR	CD2	-10.656	32.374	-3.826
171	TYR	CE2	-10.943	33.088	-1.936
171	TYR	OH	-12.008	33.119	0.170
172	PRO	CA	-9.893	26.617	-6.306
172	PRO	O	-8.325	26.704	-8.081
172	PRO	CG	-10.630	29.271	-8.896
173	SER	N	-10.097	28.167	-8.019
173	SER	C	-9.828	29.773	-9.093
173	SER	CB	-11.520	29.023	-9.481
174	VAL	N	-8.162	29.944	-8.616
174	VAL	C	-8.754	30.231	-9.060
174	VAL	CB	-6.099	31.775	-7.896
174	VAL	CG2	-8.220	32.803	-7.323
175	ILE	CA	-3.869	30.186	-10.024
175	ILE	O	-2.650	31.050	-8.933
175	ILE	CG1	-3.857	29.978	-12.924
175	ILE	CD1	-3.692	30.829	-13.066
176	ALA	CA	-1.335	30.317	-6.870
176	ALA	O	0.633	29.218	-7.038
177	VAL	N	0.064	31.410	-7.100
177	VAL	C	3.233	31.693	-6.673
177	VAL	CB	2.439	32.007	-8.703
177	VAL	CG2	1.374	32.832	-9.048
178	GLY	CA	3.160	30.703	-8.339
178	GLY	O	0.609	31.438	-7.206
179	ALA	CA	0.718	32.837	-9.099
179	ALA	C	10.100	38.601	-6.719
180	VAL	N	10.639	31.102	-6.009
180	VAL	C	11.840	31.803	-7.171
180	VAL	CB	12.878	29.816	-8.166
180	VAL	CG2	11.679	30.320	-9.900
181	ASP	CA	13.631	32.100	-7.839
181	ASP	O	13.330	31.008	-9.802
181	ASP	CG	17.120	30.936	-8.971
181	ASP	OD2	17.600	30.216	-6.007
182	SER	CA	17.622	32.816	-10.101
182	SER	O	10.365	30.692	-11.670
182	SER	OG	10.816	36.061	-10.678
183	SER	CA	10.718	30.068	-9.666
183	SER	O	17.039	26.618	-9.397

165	VAL	CG1	-1.047	30.397	-1.374
166	GLY	N	-1.918	31.021	1.129
166	GLY	C	-6.098	32.089	0.017
167	TYR	N	-8.086	33.730	0.970
167	TYR	C	-8.993	33.309	-0.606
167	TYR	CB	-7.666	36.232	0.966
167	TYR	CD1	-7.800	32.703	2.067
167	TYR	CE1	-7.367	31.920	3.618
167	TYR	CZ	-8.486	30.671	3.046
168	PRO	N	-6.380	33.490	-1.090
168	PRO	CB	-6.873	36.782	-2.624
168	PRO	CA	-7.136	36.637	-2.869
168	PRO	O	-7.097	32.828	-3.012
169	GLY	CA	-6.666	38.077	-3.927
169	GLY	O	-6.000	30.733	-6.269
170	LYS	CA	-8.096	30.260	-1.763
170	LYS	O	-7.300	27.304	-2.824
170	LYS	CG	-8.798	28.106	0.903
170	LYS	CE	-8.731	27.271	3.029
171	TYR	N	-7.030	29.016	-3.160
171	TYR	C	-8.603	28.309	-3.113
171	TYR	CB	-9.962	38.224	-6.262
171	TYR	CD1	-11.060	38.303	-1.982
171	TYR	CE1	-11.820	31.883	-0.067
171	TYR	CZ	-11.820	32.398	-0.006
172	PRO	N	-9.297	27.804	-8.376
172	PRO	C	-9.233	27.196	-7.009
172	PRO	CB	-10.167	28.329	-6.813
172	PRO	CD	-10.364	26.669	-6.816
173	SER	CA	-10.220	28.018	-9.330
173	SER	O	-8.966	38.233	-10.762
173	SER	OG	-11.895	39.846	-0.686
174	VAL	CA	-7.033	30.091	-8.095
174	VAL	O	-8.612	29.182	-8.366
174	VAL	CG1	-8.796	32.837	-7.617
175	ILE	N	-6.911	30.729	-9.005
175	ILE	C	-2.714	38.736	-8.094
175	ILE	CB	-2.933	30.826	-11.610
175	ILE	CG2	-1.681	30.009	-11.812
176	ALA	N	-2.820	30.820	-7.928
176	ALA	C	0.120	30.301	-7.310
176	ALA	CB	-1.639	29.030	-8.861
177	VAL	CA	2.861	31.834	-7.036
177	VAL	O	3.170	32.697	-8.721
177	VAL	CG1	3.062	32.667	-9.392
178	GLY	N	6.077	30.694	-6.390
178	GLY	C	6.666	31.233	-6.076
179	ALA	N	7.012	31.667	-8.207
179	ALA	C	9.939	31.099	-8.770
179	ALA	CB	9.029	33.231	-6.973
180	VAL	CA	11.070	38.603	-6.901
180	VAL	O	12.712	32.691	-7.627
180	VAL	CG1	11.271	30.291	-7.098
181	ASP	N	16.267	31.203	-6.080
181	ASP	C	18.062	31.006	-0.662
181	ASP	CB	16.666	31.921	-8.916
181	ASP	OD1	17.108	29.703	-6.072
182	SER	N	17.007	32.306	-8.047
182	SER	C	18.103	30.017	-10.696
182	SER	CB	18.070	33.313	-10.666
183	SER	N	18.280	30.062	-9.023
183	SER	C	17.801	27.616	-9.067
183	SER	CB	18.806	38.823	-8.087

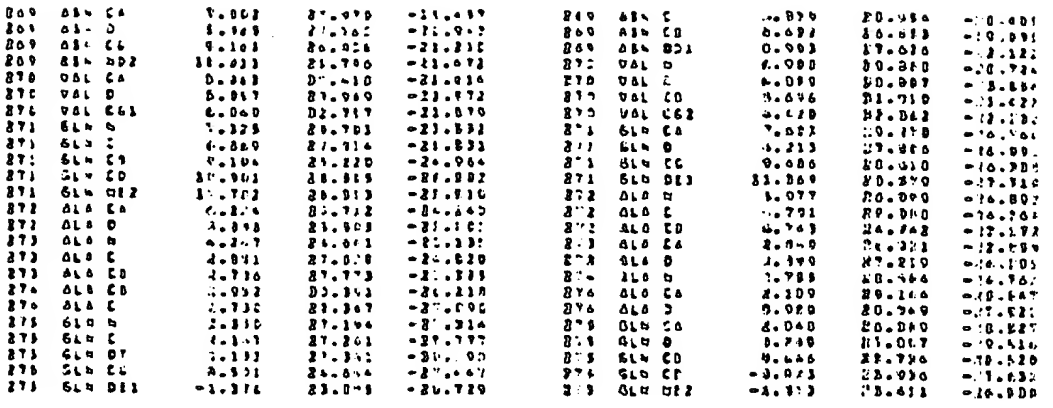
183	SER	OG	20.889	20.615	-0.251	184	ASN	N	16.973	20.094	-0.602
184	ASN	CA	19.144	27.317	-0.990	184	ASN	C	16.931	26.728	-0.197
184	ASN	O	16.178	25.759	-0.097	184	ASN	CB	18.016	26.241	-10.712
184	ASN	CG	16.993	26.998	-12.076	184	ASN	OD1	16.709	28.184	-12.277
184	ASN	ND2	15.952	26.218	-13.076	185	GLN	N	15.042	27.247	-7.199
185	GLN	CA	15.276	26.646	-5.039	185	GLN	C	14.209	27.694	-5.203
185	GLN	O	14.199	28.726	-5.396	185	GLN	CD	16.099	26.960	-5.101
185	GLN	CG	16.539	26.242	-5.614	185	GLN	CD	15.011	26.102	-5.204
185	GLN	OE1	16.864	25.799	-6.061	185	GLN	NE2	15.266	26.806	-1.934
186	ARG	N	13.270	26.958	-4.448	186	ARG	CA	12.185	27.774	-5.041
186	ARG	C	12.700	28.782	-2.066	186	ARG	O	13.098	29.304	-2.093
186	ARG	CD	11.215	26.043	-5.116	186	ARG	CG	10.214	27.671	-2.161
186	ARG	CB	0.467	26.337	-1.661	186	ARG	NE	9.066	26.933	-0.117
186	ARG	CZ	0.961	26.879	1.939	186	ARG	NH1	9.367	27.008	1.658
186	ARG	NH2	18.066	26.321	1.783	187	ALA	N	12.204	28.009	-2.093
187	ALA	CA	12.728	31.064	-1.095	187	ALA	C	12.262	28.694	-0.817
187	ALA	B	11.151	30.043	-8.307	187	ALA	CB	13.144	28.492	-2.544
188	SER	N	13.991	32.770	0.549	188	SER	CA	12.671	30.206	1.060
188	SER	C	11.356	33.047	2.412	188	SER	O	10.748	28.111	2.212
188	SER	CD	13.767	30.686	2.938	188	SER	OG	14.137	31.026	2.041
189	PHE	N	18.963	32.010	1.095	189	PHE	CA	0.697	32.000	2.610
189	PHE	C	0.499	32.198	1.609	189	PHE	O	9.309	32.956	2.011
189	PHE	CD	9.787	34.217	2.248	189	PHE	CG	10.117	34.096	0.057
189	PHE	CD1	9.167	34.930	-0.121	189	PHE	CD2	11.418	33.116	0.967
189	PHE	CE1	9.483	35.107	-1.411	189	PHE	CE2	11.760	33.948	-0.701
189	PHE	CZ	10.786	35.806	-1.725	190	SER	N	0.783	31.926	0.699
190	SER	CA	7.626	31.896	-8.301	190	SER	C	6.663	30.162	0.928
190	SER	O	7.034	29.803	0.066	190	SER	CD	0.101	28.998	-1.700
190	SER	OG	7.136	33.337	-2.618	191	SER	N	5.000	30.951	0.026
191	SER	CA	6.341	29.696	0.987	191	SER	C	4.261	30.338	0.223
191	SER	O	4.363	30.268	-6.095	191	SER	CD	3.913	28.411	0.911
191	SER	OG	2.729	31.205	1.054	192	VAL	N	3.786	27.318	0.928
192	VAL	CA	3.629	25.032	0.301	192	VAL	C	2.284	25.291	0.606
192	VAL	B	1.959	25.698	1.998	192	VAL	CD	4.701	25.127	1.008
192	VAL	CG1	6.144	25.727	0.722	192	VAL	CG2	4.617	25.196	2.092
193	GLY	N	1.938	24.172	0.847	193	GLY	CA	0.629	23.864	0.416
193	GLY	C	0.881	23.829	-8.981	193	GLY	B	0.538	23.264	-2.019
194	PRO	N	-1.823	22.209	-8.722	194	PRO	CA	-1.662	21.851	-1.873
194	PRO	C	-2.237	22.695	-2.914	194	PRO	B	-2.683	22.264	-6.005
194	PRO	CD	-2.769	20.703	-1.218	194	PRO	CG	-2.311	20.622	0.213
194	PRO	CD	-1.633	21.954	0.070	195	GLU	N	-2.022	25.703	-2.639
195	GLU	CA	-3.145	24.050	-8.252	195	GLU	C	-2.095	29.021	-6.058
195	GLU	B	-2.516	26.300	-6.936	195	GLU	CD	-6.043	29.706	-8.670
195	GLU	CG	-6.042	25.134	-1.633	195	GLU	CD	-6.919	26.068	-0.109
195	GLU	OE1	-3.110	24.960	0.165	195	GLU	OE2	-8.130	24.520	0.703
196	LEU	N	-0.529	25.264	-8.070	196	LEU	CA	0.241	28.929	-6.664
196	LEU	C	0.228	21.376	-6.059	196	LEU	B	0.305	24.121	-6.193
196	LEU	CD	1.945	25.789	-8.094	196	LEU	CG	2.770	26.178	-6.643
196	LEU	CD1	2.739	27.716	-6.630	196	LEU	CD2	4.027	29.721	-8.911
197	ASP	N	0.148	26.288	-7.993	197	ASP	CA	0.032	29.774	-8.408
197	ASP	C	1.337	28.721	-9.293	197	ASP	B	1.053	26.734	-9.916
197	ASP	CD	-1.067	26.998	-9.191	197	ASP	CG	-2.686	26.991	-8.869
197	ASP	OD1	-2.084	28.158	-8.954	197	ASP	OD2	-3.936	27.327	-9.008
198	VAL	N	2.813	26.009	-9.844	198	VAL	CA	3.286	26.978	-10.289
198	VAL	C	6.197	27.958	-9.814	198	VAL	B	3.752	28.699	-9.927
198	VAL	CD	2.894	27.476	-11.037	198	VAL	CG1	3.938	26.726	-12.937
198	VAL	CG2	3.337	28.919	-11.404	199	GLY	N	5.976	27.916	-10.016
199	GLY	CA	6.439	28.882	-9.498	199	GLY	C	6.068	29.018	-10.870
199	GLY	B	6.696	29.318	-11.793	199	GLY	CD	7.669	27.970	-9.077
199	GLY	CG	7.363	26.049	-8.139	199	GLY	OD	6.783	27.669	-6.568
199	GLY	CD	0.227	27.785	-8.587	200	ALA	N	7.626	30.942	-10.103
200	ALA	CA	7.901	31.929	-11.058	200	ALA	C	9.000	32.666	-10.872
200	ALA	B	8.127	32.924	-9.068	200	ALA	CD	6.932	32.070	-11.630

201	PRO	N	9.927	58.499	-10.983	201	PRO	CA	11.013	54.180	-10.210
201	PRO	C	10.480	55.127	-9.238	201	PRO	O	9.570	55.007	-9.612
201	PRO	CB	11.037	54.723	-11.600	201	PRO	CG	11.592	54.040	-12.670
201	PRO	CD	9.941	53.616	-12.485	202	GLY	N	10.925	55.206	-8.021
202	GLY	CA	10.473	56.234	-7.044	202	GLY	C	11.500	56.698	-6.113
202	GLY	O	11.382	57.124	-6.979	203	VAL	N	12.018	56.583	-6.613
203	VAL	CA	13.940	56.920	-5.716	203	VAL	C	14.706	55.017	-6.669
203	VAL	C	13.133	57.731	-7.803	203	VAL	CB	15.016	55.600	-5.581
203	VAL	CG1	16.006	58.106	-6.632	203	VAL	CG2	14.070	54.761	-4.570
204	SER	N	14.065	59.102	-5.039	204	SER	CA	15.572	60.201	-6.607
204	SER	C	15.067	60.619	-7.072	204	SER	C	15.706	60.605	-8.089
204	SER	CB	17.087	59.976	-6.326	204	SER	OG	17.732	61.106	-6.672
205	ILE	N	13.771	60.965	-8.000	205	ILE	CA	13.059	61.236	-9.225
205	ILE	C	13.207	62.749	-9.670	205	ILE	O	13.670	63.690	-8.040
205	ILE	CB	11.832	60.833	-9.144	205	ILE	CG1	11.636	59.536	-8.010
205	ILE	CG2	10.099	61.281	-10.667	205	ILE	CD1	12.257	58.412	-9.771
206	GLN	N	13.956	63.995	-10.409	206	GLN	CA	14.206	64.517	-10.034
206	GLN	C	13.802	64.970	-11.630	206	GLN	O	12.669	64.510	-12.621
206	GLN	CB	15.435	64.700	-11.740	206	GLN	CG	16.606	66.163	-10.980
206	GLN	CD	17.205	65.165	-10.807	206	GLN	OE1	18.320	64.036	-9.353
206	GLN	NE2	16.336	66.260	-9.057	207	SER	N	12.399	65.064	-11.214
207	SER	CA	11.217	66.371	-11.907	207	SER	C	11.009	60.803	-11.769
207	SER	O	11.910	68.657	-11.004	207	SER	CB	9.910	69.053	-11.069
207	SER	OG	8.993	66.836	-12.613	208	THR	N	10.084	68.604	-12.326
208	THR	CG2	9.171	50.339	-14.754	208	THR	OG1	7.570	69.414	-13.164
208	THR	CB	8.620	50.419	-13.357	208	THR	CA	9.679	50.092	-12.173
208	THR	C	9.107	50.480	-10.003	208	THR	O	8.623	69.087	-10.069
209	LEU	N	9.650	51.610	-10.220	209	LEU	CA	9.192	52.150	-8.959
209	LEU	C	8.673	53.610	-9.262	209	LEU	O	9.160	54.827	-10.222
209	LEU	CB	10.335	52.192	-7.088	209	LEU	CG	10.084	55.016	-7.616
209	LEU	CD1	11.960	51.114	-6.472	209	LEU	CD2	9.607	55.202	-6.069
210	PRO	N	7.790	54.139	-8.444	210	PRO	CA	7.273	55.517	-8.049
210	PRO	C	8.303	56.573	-8.639	210	PRO	O	9.401	56.665	-8.104
210	PRO	CB	6.382	55.733	-7.817	210	PRO	CG	6.884	54.370	-6.964
210	PRO	CD	7.193	53.491	-7.271	211	GLY	N	8.977	57.065	-9.335
211	GLY	CA	9.069	58.765	-9.610	211	GLY	C	10.896	55.696	-10.698
211	GLY	O	11.176	59.008	-10.289	212	ASN	N	9.091	57.770	-11.007
212	ASN	CA	10.983	57.422	-12.643	212	ASN	C	12.059	56.793	-12.096
212	ASN	C	13.100	57.101	-12.620	212	ASN	CB	11.224	55.595	-13.699
212	ASN	CG	11.083	50.109	-14.016	212	ASN	OD1	11.093	57.084	-15.323
212	ASN	ND2	12.273	59.159	-15.376	213	LYS	N	11.003	55.769	-11.847
213	LYS	CA	13.010	54.040	-10.937	213	LYS	C	12.000	53.659	-10.066
213	LYS	O	11.778	53.039	-11.613	213	LYS	CB	12.709	55.241	-9.059
213	LYS	CG	13.286	56.094	-8.707	213	LYS	CD	13.266	57.050	-7.312
213	LYS	CE	14.135	58.210	-6.070	213	LYS	NZ	15.040	55.785	-7.021
214	TYR	N	13.601	52.703	-10.444	214	TYR	CA	13.003	51.246	-10.722
214	TYR	C	16.383	50.600	-9.609	214	TYR	O	15.211	51.293	-8.017
214	TYR	CB	14.641	50.981	-11.904	214	TYR	CG	14.130	51.621	-13.266
214	TYR	CD1	14.609	52.047	-13.670	214	TYR	CD2	13.129	51.065	-14.016
214	TYR	CE1	16.230	53.473	-14.016	214	TYR	CE2	12.094	51.609	-15.170
214	TYR	CZ	13.204	52.093	-15.050	214	TYR	OH	12.756	53.650	-16.696
215	GLY	N	16.080	60.047	-9.190	215	GLY	CA	16.622	60.772	-7.905
215	GLY	C	16.138	67.320	-7.769	215	GLY	O	13.209	66.917	-8.521
216	ALA	N	16.010	66.690	-8.031	216	ALA	CA	16.684	65.809	-6.701
216	ALA	C	15.602	64.922	-8.912	216	ALA	O	13.960	65.327	-6.479
216	ALA	CB	15.718	64.354	-6.007	217	TYR	N	12.750	63.082	-5.979
217	TYR	CA	11.064	63.600	-6.646	217	TYR	C	12.033	61.920	-6.547
217	TYR	O	12.262	61.642	-5.636	217	TYR	CB	10.673	65.062	-6.570
217	TYR	CG	10.117	60.291	-4.216	217	TYR	CD1	10.046	65.991	-3.236
217	TYR	CD2	9.516	65.933	-6.705	217	TYR	CE1	10.650	67.807	-2.790
217	TYR	CE2	8.654	67.210	-6.301	217	TYR	CZ	9.550	67.002	-3.501
217	TYR	OH	8.933	69.160	-2.900	218	ASN	N	11.708	61.550	-3.591
218	ASN	CA	11.646	59.042	-3.227	218	ASN	C	10.204	59.036	-2.769

210	ASN	B	0.763	53.047	-1.917	218	ASN	CB	12.953	39.340	-3.194
210	ASN	CG	14.031	39.066	-2.342	218	ASN	OD1	14.612	39.789	-3.472
210	ASN	ND2	14.668	39.044	-1.168	219	GLY	N	9.670	38.934	-2.289
219	GLY	CA	0.302	38.138	-2.649	219	GLY	C	7.870	37.304	-1.681
219	GLY	O	7.078	37.802	-4.076	220	THR	N	6.561	34.638	-3.205
220	THR	CA	5.697	35.936	-4.179	220	THR	C	4.079	37.044	-4.864
220	THR	B	6.617	36.742	-5.018	220	THR	CB	4.825	34.819	-3.926
220	THR	OG1	6.196	38.043	-2.491	220	THR	CG2	3.784	33.696	-2.980
221	SER	N	6.730	38.238	-6.303	221	SER	CA	3.984	39.201	-5.169
221	SER	C	6.760	39.641	-6.303	221	SER	O	6.117	40.298	-7.277
221	SER	CB	3.329	40.383	-6.546	221	SER	OG	3.435	40.282	-3.140
222	MET	N	6.865	39.389	-5.605	222	MET	CB	6.471	42.771	-5.173
222	MET	SD	7.760	41.833	-6.993	222	MET	CG	8.906	41.399	-6.602
222	MET	CE	8.861	42.818	-7.210	222	MET	CA	6.916	39.678	-7.638
222	MET	C	6.077	38.635	-8.867	222	MET	O	7.804	38.567	-9.778
223	ALA	N	6.856	37.266	-8.841	223	ALA	CB	6.669	36.020	-8.885
223	ALA	C	8.200	36.868	-9.707	223	ALA	O	8.153	35.940	-10.929
223	ALA	CB	6.989	36.887	-7.923	224	SER	N	6.876	36.360	-9.838
224	SER	CA	8.798	36.488	-9.703	224	SER	C	8.661	37.161	-11.039
224	SER	B	8.165	36.893	-11.057	224	SER	CB	1.001	36.991	-8.603
224	SER	OG	0.492	36.099	-9.197	225	PRO	N	8.156	38.411	-11.199
225	PRO	CA	8.095	39.130	-12.439	225	PRO	C	9.764	38.669	-13.626
225	PRO	B	8.486	38.650	-14.004	225	PRO	CB	8.693	40.811	-12.084
225	PRO	CG	6.611	40.402	-10.764	225	PRO	CD	8.735	39.224	-10.094
226	HIS	N	6.709	37.626	-13.299	226	HIS	CA	9.646	36.879	-14.062
226	HIS	C	6.610	38.947	-15.061	226	HIS	O	6.425	38.059	-16.293
226	HIS	CB	6.088	36.946	-13.765	226	HIS	CG	7.814	36.899	-13.388
226	HIS	ND1	8.060	37.608	-12.170	226	HIS	CD2	8.983	37.110	-14.167
226	HIS	CE1	9.279	38.052	-12.236	226	HIS	NE2	9.771	37.066	-13.663
227	VAL	N	5.503	35.366	-14.199	227	VAL	CA	2.083	34.300	-14.727
227	VAL	C	3.479	35.197	-15.621	227	VAL	O	1.010	34.773	-16.695
227	VAL	CB	2.183	33.644	-13.619	227	VAL	CG1	1.976	32.476	-14.246
227	VAL	CG2	2.204	32.669	-12.891	228	ALA	N	1.903	36.242	-14.814
228	ALA	CA	0.811	37.389	-15.517	228	ALA	C	0.843	37.938	-16.068
228	ALA	B	-0.253	37.683	-17.020	228	ALA	CB	-0.307	38.383	-14.668
229	GLY	N	1.791	38.928	-16.941	229	GLY	CA	2.352	38.608	-18.239
229	GLY	C	2.620	37.197	-19.187	229	GLY	O	2.109	37.375	-20.304
230	ALA	N	3.711	38.988	-18.646	230	ALA	CA	2.794	34.801	-19.046
230	ALA	C	3.424	36.880	-20.153	230	ALA	O	1.300	34.265	-21.343
230	ALA	CB	3.290	33.624	-18.789	231	ALA	N	0.303	36.623	-19.328
231	ALA	CA	-1.018	36.416	-19.744	231	ALA	C	-1.256	38.423	-20.064
231	ALA	O	-1.089	38.858	-21.052	231	ALA	CB	-1.932	36.664	-18.949
232	ALA	N	-0.770	36.687	-21.721	232	ALA	CA	-1.013	37.063	-21.702
232	ALA	C	-0.281	37.204	-23.070	232	ALA	O	-0.041	37.981	-24.107
232	ALA	CB	-0.742	39.121	-21.377	233	LEU	N	0.938	36.726	-22.067
233	LEU	CA	1.617	36.293	-24.209	233	LEU	C	0.821	38.169	-26.006
233	LEU	B	0.696	35.231	-26.111	233	LEU	CB	2.063	35.077	-25.967
233	LEU	CG	3.996	36.994	-23.693	233	LEU	CD1	5.299	36.362	-22.921
233	LEU	CD2	4.241	37.093	-24.600	234	ILE	N	0.337	36.199	-24.947
234	ILE	CD1	0.300	38.604	-21.657	234	ILE	CG1	0.494	31.223	-23.183
234	ILE	CD	-0.011	32.814	-23.570	234	ILE	CG2	-1.803	36.980	-24.091
234	ILE	CA	-0.606	33.976	-24.646	234	ILE	C	-1.621	33.997	-25.634
234	ILE	B	-1.083	33.144	-26.844	235	LEU	N	-2.396	34.665	-24.779
235	LEU	CA	-3.896	35.828	-25.623	235	LEU	C	-3.258	39.063	-26.672
235	LEU	B	-4.189	35.914	-27.909	235	LEU	CB	-4.632	35.765	-24.378
235	LEU	CG	-5.149	36.990	-23.342	235	LEU	CD1	-5.652	39.603	-22.169
235	LEU	CD2	-6.282	36.130	-24.120	236	SER	N	-2.994	36.438	-26.700
236	SER	CA	-1.764	37.237	-27.986	236	SER	C	-1.691	36.292	-29.146
236	SER	B	-1.766	36.636	-29.290	236	SER	CB	-0.633	38.234	-27.733
236	SER	OG	0.090	37.071	-27.982	237	LYS	N	-1.846	35.867	-28.082
237	LYS	CA	-0.060	36.885	-29.952	237	LYS	C	-2.113	33.277	-30.260
237	LYS	B	-2.370	32.981	-31.444	237	LYS	CB	0.272	33.112	-29.551
237	LYS	CE	0.677	32.260	-30.716	237	LYS	CD	2.020	31.035	-30.447

DD7	LYS	CE	8.308	D0.702	-D1.720		837	LYS	OT	D.020	80.040	-81.806
DD0	α13	α	-8.081	D1.909	-29.312		830	α18	CD	-6.800	D2.103	-20.170
DD0	α11	C	-8.D36	D2.899	-80.697		8D0	α11	D	-D.718	D2.984	-27.862
890	α18	CD	-D.040	D8.062	-28.811		8D0	α11	CE	-D.000	80.021	-20.237
DD0	α18	OE1	-3.787	89.670	-28.888		8D0	α18	CD2	-D.337	80.200	-38.384
8D0	α13	CE1	-1.006	80.081	-89.048		8D0	α18	OE2	-1.060	80.008	-38.899
8D9	PDD	α	-D.040	D3.917	-80.208		8D0	PDD	CD	-0.000	D6.770	-80.773
839	PDD	C	-0.286	D6.882	-20.882		8D0	PDD	D	-0.969	D6.019	-87.062
839	PDD	CD	-7.010	D8.977	-20.713		8D9	PDD	CE	-6.666	D9.806	-D1.027
8D9	PDD	CD	-D.636	D6.630	-3D.660		868	ASN	α	-D.D06	D2.069	-20.227
860	ASN	CA	-9.029	D2.961	-89.216		860	ASN	C	-0.990	D1.100	-87.008
860	ASN	D	-10.040	D0.610	-87.876		860	ASN	CD	-9.683	D1.269	-38.838
860	ASN	CE	-7.071	DE.027	-DD.009		860	ASN	OD1	-7.098	D1.900	-31.167
860	ASN	OD2	-7.67E	89.889	-DE.986		861	YDP	α	-0.D86	D1.086	-87.D06
861	YDP	CA	-8.D06	D8.126	-26.12E		861	YDP	C	-0.106	D8.630	-26.036
861	YDP	D	-9.063	D1.033	-26.606		861	YDP	CD	-6.079	89.03E	-28.670
861	YDP	CE	-6.096	8D.983	-26.887		861	YDP	CD1	-6.380	80.033	-27.810
861	YDP	CD2	-6.030	88.826	-26.188		861	YDP	OE1	-D.862	87.867	-30.211
861	YDP	CE2	-6.616	87.676	-27.216		861	YDP	CE3	-6.897	80.606	-26.981
861	YDP	CE2	-D.109	86.706	-27.176		861	YDP	CE3	-2.012	87.067	-26.043
861	YDP	CH2	-2.670	86.073	-26.008		862	YAO	α	-0.727	89.701	-26.162
862	YAO	CA	-10.680	DE.119	-22.911		862	YAO	C	-0.669	D0.176	-21.767
862	YAO	D	-8.833	89.676	-21.937		862	YAO	CD	-11.070	89.032	-22.678
862	YAO	OG1	-10.037	87.706	-22.676		862	YAO	CG2	-12.606	80.087	-23.008
863	ASN	α	-0.060	89.089	-20.611		863	ASN	OD2	-11.707	D8.606	-10.767
863	ASN	OD1	-11.668	D1.818	-16.780		863	ASN	CE	-11.003	D1.391	-17.908
863	ASN	CD	-9.700	D1.830	-10.332		863	ASN	CA	-0.098	D0.731	-10.666
863	ASN	C	-8.087	89.363	-19.010		863	ASN	D	-7.093	D0.196	-10.669
866	YAO	D	-9.066	88.362	-19.809		866	YAO	CA	-0.301	86.986	-10.080
866	YAO	C	-0.133	86.393	-19.082		866	YAO	D	-7.826	8D.707	-19.311
866	YAO	CD	-10.068	86.800	-10.696		866	YAO	OG1	-11.790	86.678	-10.686
866	YAO	CG2	-12.933	86.998	-19.189		868	GLN	α	-D.E02	86.716	-21.073
868	GLN	CA	-6.966	86.362	-21.062		868	GLN	C	-D.067	87.020	-21.920
868	GLN	D	-6.873	86.393	-21.667		868	GLN	CD	-7.D30	86.099	-23.397
868	GLN	CE	-0.268	88.826	-23.900		868	GLN	CD	-0.693	80.073	-20.620
868	GLN	OE1	-9.306	86.769	-23.727		868	GLN	NE2	-7.960	89.312	-26.370
866	VAL	α	-8.697	88.306	-21.210		866	VAL	CA	-6.677	80.060	-28.770
866	VAL	C	-D.036	86.662	-10.667		866	VAL	D	-2.780	80.227	-10.961
866	VAL	CD	-6.779	D0.888	-28.621		866	VAL	CG1	-3.066	D3.272	-28.027
866	VAL	CG2	-D.169	D1.330	-21.980		867	ARG	α	-6.707	80.260	-10.662
867	ARG	CA	-6.30E	87.716	-17.160		867	ARG	C	-D.770	86.292	-17.960
867	ARG	D	-2.700	88.008	-16.706		867	ARG	CD	-D.033	87.667	-16.169
867	ARG	CE	-6.907	87.898	-16.082		867	ARG	CD	-6.096	87.170	-1D.703
867	ARG	NE	-D.640	86.787	-12.060		867	ARG	CZ	-D.093	80.066	-11.318
867	ARG	OH1	-7.806	87.606	-11.210		867	ARG	OH2	-D.377	80.620	-10.270
860	SER	α	-6.600	88.809	-10.131		860	SER	CA	-6.030	86.331	-10.626
860	SER	C	-2.087	86.906	-10.072		860	SER	D	-1.060	83.393	-10.903
860	SER	CD	-D.836	83.690	-10.372		860	SER	OG	-6.166	83.00E	-10.832
869	SER	α	-2.900	86.033	-20.130		869	SER	CD	-1.223	86.076	-20.081
869	SER	C	-8.071	8D.302	-10.068		869	SER	D	8.026	86.700	-20.069
869	SER	CD	-1.060	8D.780	-22.060		869	SER	OG	-D.308	8D.619	-22.986
880	LEU	α	-0.209	86.333	-10.169		880	LEU	CD2	3.026	80.036	-10.222
880	LEU	CD1	-0.370	D8.633	-17.260		880	LEU	CE	0.382	80.680	-10.101
880	LEU	CD	0.170	80.061	-17.008		880	LEU	CA	0.780	80.087	-10.216
880	LEU	C	3.892	80.094	-17.208		880	LEU	C	8.303	8D.621	-17.032
881	GLN	α	0.900	8D.0E7	-16.716		881	GLN	OE2	-8.708	8D.032	-12.237
881	GLN	OE1	-8.019	83.626	-12.930		881	GLN	CD	-8.D60	86.080	-13.036
881	GLN	CE	-3.210	86.016	-13.996		881	GLN	CD	-0.007	8D.621	-16.077
881	GLN	CA	0.D01	23.061	-13.760		881	GLN	C	0.909	82.666	-10.D61
881	GLN	D	8.768	82.816	-19.010		882	ASN	α	0.0DD	82.386	-17.D0E
882	ASN	CA	8.002	81.206	-10.802		882	ASN	C	8.D06	81.300	-10.991
882	ASN	D	8.089	80.662	-19.760		882	ASN	CD	0.006	80.700	-19.202
882	ASN	CE	-3.836	19.026	-10.078		882	ASN	OD1	-0.0D6	19.008	-17.902

851	ASN	OD2	-2.894	19.024	-19.161	883	TYR	N	8.018	22.583	-18.921
852	TYR	CA	6.256	22.717	-19.713	883	TYR	C	9.301	23.247	-18.818
853	TYR	O	6.348	23.733	-19.427	883	TYR	CB	6.004	23.672	-20.982
853	TYR	CB1	5.893	24.057	-20.428	883	TYR	CE2	8.147	23.138	-22.032
854	TYR	N	8.218	23.177	-17.951	884	TYR	CA	9.216	23.612	-16.988
854	TYR	C	9.466	22.730	-16.612	884	TYR	O	9.402	23.900	-17.093
854	TYR	CD	8.664	23.998	-15.132	884	TYR	CB1	8.129	22.170	-15.048
854	TYR	CE2	6.830	24.849	-14.882	885	TYR	N	8.499	23.296	-16.076
855	TYR	CA	8.771	22.894	-13.817	885	TYR	C	9.621	22.931	-14.414
855	TYR	O	9.439	22.786	-13.474	885	TYR	CD	11.883	23.483	-15.097
855	TYR	CB1	11.082	23.709	-17.921	885	TYR	CE2	12.206	22.628	-15.496
856	LYS	N	9.606	25.702	-14.314	886	LYS	CA	9.364	20.063	-13.015
856	LYS	C	10.322	20.333	-12.063	886	LYS	O	11.662	20.274	-12.992
856	LYS	CD	9.024	18.698	-13.249	886	LYS	CG	9.818	17.005	-11.921
856	LYS	CD	10.206	16.948	-11.777	886	LYS	CE	10.232	19.948	-10.623
856	LYS	NZ	9.243	16.969	-11.854	887	LEU	N	10.212	20.674	-10.824
887	LEU	CA	11.272	21.038	-9.893	887	LEU	C	11.293	20.232	-8.614
887	LEU	O	12.096	20.365	-7.732	887	LEU	CB	11.187	22.547	-9.822
887	LEU	CG	11.357	23.620	-10.068	887	LEU	CD1	11.248	25.093	-9.921
887	LEU	CD2	12.678	23.468	-11.325	888	GLY	N	10.431	19.282	-8.298
888	GLY	CA	10.602	18.793	-6.879	888	GLY	C	9.108	18.703	-6.373
888	GLY	O	8.283	18.956	-7.282	889	ASP	N	9.024	18.282	-5.190
889	ASP	CA	7.757	17.896	-4.816	889	ASP	C	6.699	18.941	-6.709
889	ASP	O	6.099	20.039	-4.214	889	ASP	CB	7.996	17.840	-3.083
889	ASP	CG	6.781	17.128	-2.841	889	ASP	OD1	8.611	17.827	-2.354
889	ASP	OD2	7.095	16.299	-1.321	890	SER	N	5.560	18.618	-5.312
890	SER	CA	6.681	19.587	-5.529	890	SER	C	6.046	20.362	-4.209
890	SER	O	5.500	21.503	-4.646	890	SER	CB	5.348	18.919	-6.289
890	SER	OG	5.745	17.937	-5.648	891	PHE	N	6.241	19.770	-3.112
891	PHE	CA	5.831	20.668	-1.865	891	PHE	C	6.544	21.046	-1.863
891	PHE	O	5.944	22.060	-1.632	891	PHE	CD	6.053	19.749	-0.563
891	PHE	CG	5.549	20.337	0.719	891	PHE	CD1	2.206	20.103	1.125
891	PHE	CD2	6.691	21.060	1.938	891	PHE	CE1	3.737	20.717	2.315
891	PHE	CE2	3.949	21.602	2.748	891	PHE	CZ	3.605	21.469	3.114
892	TYR	N	5.778	21.788	-2.305	892	TYR	CA	6.680	22.914	-2.291
892	TYR	C	6.822	23.689	-3.949	892	TYR	O	7.781	24.053	-3.293
892	TYR	CD	8.122	22.493	-1.051	892	TYR	CG	8.146	21.092	-0.454
892	TYR	CD1	9.984	20.484	-0.364	892	TYR	CD2	8.149	22.648	0.698
892	TYR	CE1	8.862	19.873	0.082	892	TYR	CE2	8.114	22.969	1.962
892	TYR	CZ	8.869	20.672	2.910	892	TYR	OH	7.965	20.929	3.205
893	TYR	N	6.626	23.184	-4.693	893	TYR	CA	6.012	23.658	-6.022
893	TYR	C	8.626	23.600	-6.956	893	TYR	O	8.781	24.117	-8.111
893	TYR	CB	7.020	22.768	-6.681	893	TYR	CG	9.279	23.035	-6.068
893	TYR	CD1	10.864	24.044	-6.637	893	TYR	CD2	9.005	22.942	-4.998
893	TYR	CE1	11.338	24.328	-8.168	893	TYR	CE2	11.862	22.668	-4.491
893	TYR	CZ	11.038	23.618	-8.186	893	TYR	OH	13.865	23.949	-4.097
894	GLY	N	6.671	23.161	-6.010	894	GLY	CA	8.581	23.864	-7.612
894	GLY	C	8.047	22.196	-8.856	894	GLY	O	6.047	21.274	-8.865
895	LYS	N	8.636	22.477	-9.794	895	LYS	CA	9.034	21.798	-10.971
895	LYS	C	8.108	21.232	-11.664	895	LYS	O	8.084	21.063	-12.806
895	LYS	CD	8.758	22.671	-12.844	895	LYS	CG	8.695	21.963	-11.008
895	LYS	CD	8.710	20.548	-12.879	895	LYS	CE	-8.692	20.696	-11.091
895	LYS	NZ	-8.070	20.797	-12.409	896	GLY	N	8.707	23.226	-10.017
896	GLY	CA	7.120	23.612	-11.923	896	GLY	C	7.198	25.052	-11.010
896	GLY	O	6.177	25.793	-11.648	897	LEU	N	8.262	25.936	-12.688
897	LEU	CA	8.698	26.868	-13.997	897	LEU	C	7.004	26.771	-14.437
897	LEU	O	7.953	25.909	-15.298	897	LEU	CB	10.010	26.055	-13.216
897	LEU	CG	10.632	28.668	-14.050	897	LEU	CD1	10.096	29.331	-13.295
897	LEU	CD2	11.024	27.921	-16.327	898	ILE	N	7.864	27.863	-14.632
898	ILE	CA	6.406	28.235	-15.044	898	ILE	C	7.436	28.246	-17.063
898	ILE	O	8.939	28.793	-16.012	898	ILE	CD	8.069	29.218	-15.099
898	ILE	CG1	6.099	28.541	-15.992	898	ILE	CG2	6.843	28.923	-14.067
898	ILE	CD1	8.399	31.765	-16.262	899	ASN	N	7.097	27.063	-18.237



The above structural studies together with the kinetic data presented herein and elsewhere (Philipp, M., et al. (1983) Mol. Cell. Biochem. 51, 5-32; Svendsen, I.B. (1976) Carlsberg Res. Comm. 41, 237-291; Markland, S.F. Id; Stauffe, D.C., et al. (1965) J. Biol. Chem. 244, 5333-5338) indicate that the subsites in the binding cleft of subtilisin are capable of interacting with substrate amino acid residues from P-4 to P-2'.

The most extensively studied of the above residues are Gly166, Gly169 and Ala152. These amino acids were identified as residues within the S-1 subsite. As seen in Fig. 3, which is a stereoview of the S-1 subsite, Gly166 and Gly169 occupy positions at the bottom of the S-1 subsite, whereas Ala152 occupies a position near the top of S-1, close to the catalytic Ser221.

All 19 amino acid substitutions of Gly166 and Gly169 have been made. As will be indicated in the examples which follow, the preferred replacement amino acids for Gly166 and/or Gly169 will depend on the specific amino acid occupying the P-1 position of a given substrate.

The only substitutions of Ala152 presently made and analyzed comprise the replacement of Ala152 with Gly and Ser. The results of these substitutions on P-1 specificity will be presented in the examples.

In addition to those residues specifically associated with specificity for the P-1 substrate amino acid, Tyr104 has been identified as being involved with P-4 specificity. Substitutions at Phe189 and Tyr217, however, are expected to respectively effect P-2' and P-1' specificity.

The catalytic activity of subtilisin has also been modified by single amino acid substitutions at Asn155. The catalytic triad of subtilisin is shown in Fig. 4. As can be seen, Ser221, His64 and Asp32 are positioned to facilitate nucleophilic attach by the serine hydoxylate on the carbonyl of the scissile peptide bond. Crystallographic studies of subtilisin (Robertus, et al. (1972) Biochem. 11, 4293-4303; Matthews, et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, et al. (1976) J. Biol. Chem. 250, 1097-1103) show that two hydrogen bonds are formed with the oxyanion of the substrate transition state. One hydrogen bond donor is from the catalytic serine-221 main-chain amide while the other is from one of the NE2 protons of the asparagine-155 side chain. See Fig. 4.

Asn155 was substituted with Ala, Asp, His, Glu and Thr. These substitutions were made to investigate the the stabilization of the charged tetrahedral intermediate of the transition state complex by the potential hydrogen bond between the side chain of Asn155 and the oxyanion of the intermediate. These particular substitutions caused large decreases in substrate turnover, kcat (200 to 4,000 fold), marginal decreases in substrate binding Km (up to 7 fold), and a loss in transition state stabilization energy of 2.2 to 4.7 kcal/mol. The retention of Km and the drop in kcat will make these mutant enzymes useful as binding proteins for specific; peptide sequences, the nature of which will be determined by the specificity of the precursor protease.

Various other amino acid residues have been identified which affect alkaline stability. In some cases, mutants having altered alkaline stability also have altered thermal stability.

In B amyloliquefaciens subtilisin residues Asp36, Ile107, Lys170, Ser204 and Lys213 have been identified as residues which upon substitution with a different amino acid alter the alkaline stability of the mutated enzyme as compared to the precursor enzyme. The substitution of Asp36 with Ala and the substitution of Lys170 with Glu each resulted in a mutant enzyme having a lower alkaline stability as compared to the wild type subtilisin. When Ile107 was substituted with Val, Ser204 substituted with Cys, Arg or Leu or Lys213 substituted with Arg, the mutant subtilisin had a greater alkaline stability as compared

to the wild type subtilisin. However, the mutant Ser204P demonstrated a decrease in alkaline stability.

In addition, other residues, identified as being associated with the modification of other properties of subtilisin, also affect alkaline stability. These residues include Ser24, Met50, Glu156, Gly166, Gly169 and Tyr217. Specifically the following particular substitutions result in an increased alkaline stability: Ser24C, Met50F, Gly156Q or S, Gly166A, H, K, N or Q, Gly169S or A, and Tyr217F, K, R or L. The mutant Met50V, on the other hand, results in a decrease in the alkaline stability of the mutant subtilisin as compared to wild type subtilisin.

Other residues involved in alkaline stability based on the alkaline stability screen include Asp197 and Met222. Particular mutants include Asp197(R or A) and Met 222 (all other amino acids).

Various other residues have been identified as being involved in thermal stability as determined by the thermal stability screen herein. These residues include the above identified residues which effect alkaline stability and Met199 and Tyr21. These latter two residues are also believed to be important for alkaline stability. Mutants at these residues include I199 and F21.

The amino acid sequence of B. amyloliquefaciens substilisin has also been modified by substituting two or more amino acids of the wild-type sequence. Six categories of multiply substituted mutant subtilisin have been identified. The first two categories comprise thermally and oxidatively stable mutants. The next three other categories comprise mutants which combine the useful properties of any of several single mutations of B. amyloliquefaciens subtilisin. The last category comprises mutants which have modified alkaline and/or thermal stability.

The first category comprises double mutants in which two cysteine residues have been substituted at various amino acid residue positions within the subtilisin molecule. Formation of disulfide bridges between the two substituted cysteine residues results in mutant subtilisins with altered thermal stability and catalytic activity. These mutants include A21/C22/C87 and C24/C87 which will be described in more detail in Example 11.

The second category of multiple subtilisin mutants comprises mutants which are stable in the presence of various oxidizing agents such as hydrogen peroxide or peracids. Examples 1 and 2 describe these mutants which include F50/I124/Q222, F50/I124, F50/Q222, F50/L124/Q222, I124/Q222 and L124/Q222.

The third category of multiple subtilisin mutants comprises mutants with substitutions at position 222 combined with various substitutions at positions 166 or 169. These mutants, for example, combine the property of oxidative stability of the A222 mutation with the altered substrate specificity of the various 166 or 169 substitutions. Such multiple mutants include A166/A222, A166/C222, F166/C222, K166/A222, K166/C222, V166/A222 and V166/C222. The K166/A222 mutant subtilisin, for example, has a kcat/Km ratio which is approximately two times greater than that of the single A222 mutant subtilisin when compared using a substrate with phenylalanine as the P-1 amino acid. This category of multiple mutant is described in more detail in Example 12.

The fourth category of multiple mutants combines substitutions at position 156 (Glu to Q or S) with the substitution of Lys at position 166. Either of these single mutations improve enzyme performance upon substrates with glutamate as the P-1 amino acid. When these single mutations are combined, the resulting multiple enzyme mutants perform better than either precursor. See Example 9.

The fifth category of multiple mutants contain the substitution of up to four amino acids of the B. amyloliquefaciens subtilisin sequence. These mutants have specific properties which are virtually identicle to the properties of the subtilisin from B. licheniformis. The subtilisin from B. licheniformis differs from B. amyloliquefaciens subtilisin at 87 out of 275 amino acids. The multiple mutant F50/S156/A169/L217 was found to have similar substrate specificity and kinetics to the licheniformis enzyme. (See Example 13.) However, this is probably due to only three of the mutations (S156, A169 and L217) which are present in the substrate binding region of the enzyme. It is quite surprising that, by making only three changes out of the 87 different amino acids between the sequence of the two enzymes, the B. amyloliquifaciens enzyme was converted into an enzyme with properties similar to B. licheniformis enzyme. Other enzymes in this series include F50/Q156/N166/L217 and F50/S156/L217.

The sixth category of multiple mutants includes the combination of substitutions at position 107 (Ile to V) with the substitution of Lys at position 213 with Arg, and the combination of substitutions of position 204 (preferably Ser to C or L but also to all other amino acids) with the substituion of Lys at position 213 with R. Other multiple mutants which have altered alkaline stability include Q156/K166, Q156/N166, S156/K166, S156/N166 (previously identified as having altered substrate specificity), and F50/S156/A169/L217 (previously identified as a mutant of B. amyloliquifaciens subtilisin having properties similar to subtilisin from B. licheniformis). The mutant F50/V107/R213 was constructed based on the observed increase in alkaline stability for the single mutants F50, V107 and R213. It was determined that the V107/R213 mutant had an increased alkaline stability as compared to the wild type subtilisin. In this particular mutant, the increased

alkaline stability was the result of the cumulative stability of each of the individual mutations. Similarly, the mutant F50/V107/R213 had an even greater alkaline stability as compared to the V107/R213 mutant indicating that the increase in the alkaline stability due to the F50 mutation was also cumulative.

Table IV summarizes the multiple mutants which have been made including those not mentioned above. In addition, based in part on the above results, substitution at the following residues in subtilisin is expected to produce a multiple mutant having increased thermal and alkaline stability: Ser24, Met50, Ile107, Glu156, Gly166, Gly169, Ser204, Lys213, Gly215, and Tyr217.

TABLE IV

Double Mutants	Triple, Quadruple or Other Multiple
C22/C87 C24/C87 V45/V48 C49/C94 C49/C95 C50/C95 C50/C110 F50/I124 F50/Q222 I124/Q222 Q156/D166 Q156/K166 Q156/N166	F50/I124/Q222 F50/L124/Q222 F50/L124/A222 A21/C22/C87 F50/S156/N166/L217 F50/Q156/N166/L217 F50/S156/A169/L217 F50/S156/L217 F50/Q156/K166/L217 F50/S156/K166/L217 F50/Q156/K166/K217 F50/S156/K166/K217 F50/V107/R213
S156/D166 S156/K166	[S153/S156/A158/G159/S160/Δ161-164/I165/S166/A169/R170]
S156/N166	L204/R213
S156/A169 A166/A222 A166/C222	R213/204A, E, Q, D, N, G, K, V, R, T, P, I, M, F, Y, W or H
F166/A222 F166/C222 K166/A222 K166/C222 V166/A222 V166/C222 A169/A222 A169/A222 A169/C222 A21/C22	V107/R213

In addition to the above identified amino acid residues, other amino acid residues of subtilisin are also considered to be important with regard to substrate specificity. Mutation of each of these residues is expected to produce changes in the substrate specificity of subtilisin. Moreover, multiple mutations among these residues and among the previously identified residues are also expected to produce subtilisin mutants having novel substrate specificity.

Particularly important residues are His67, Ile107, Leu126 and Leu135. Mutation of His67 should alter the S-1' subsite, thereby altering the specificity of the mutant for the P-1' substrate residue. Changes at this position could also affect the pH activity profile of the mutant. This residue was identified based on the inventor's substrate modeling from product inhibitor complexes.

Ile107 is involved in P-4 binding. Mutation at this position thus should alter specificity for the P-4 substrate residue in addition to the observed effect on alkaline stability. Ile107 was also identified by molecular modeling from product inhibitor complexes.

The S-2 binding site includes the Leu126 residue. Modification at this position should therefore affect P-2 specificity. Moreover, this residue is believed to be important to convert subtilisin to an amino peptidase.

The pH activity profile should also be modified by appropriate substitution. These residues were identified from inspection of the refined model, the three dimensional structure from modeling studies. A longer side chain is expected to preclude binding of any side chain at the S-2 subsite. Therefore, binding would be restricted to subsites S-1, S-1', S-2', S-3' and cleavage would be forced to occur after the amino terminal peptide.

Leu135 is in the S-4 subsite and if mutated should alter substrate specificity for P-4 if mutated. This residue was identified by inspection of the three-dimensional structure and modeling based on the product inhibitor complex of F222.

In addition to theses sites, specific amino acid residues within the segments 97-103, 126-129 and 213-215 are also believed to be important to substrate binding.

Segments 97-103 and 126-129 form an antiparallel beta sheet with the main chain of substrate residues P-4 through P-2. Mutating residues in those regions should affect the substrate orientation through main chain (enzyme) - main chain (substrate) interactions, since the main chain of these substrate residues do not interact with these particular residues within the S-4 through S-2 subsites.

Within the segment 97-103, Gly97 and Asp99 may be mutated to alter the position of residues 101-103 within the segment. Changes at these sites must be compatible, however. In B. amyloliquifaciens subtilisin Asp99 stabilizes a turn in the main chain tertiary folding that affects the direction of residues 101-103. B. licheniformis subtilisin Asp97, functions in an analogous manner.

In addition to Gly97 and Asp99, Ser101 interacts with Asp99 in B. amyliquefaciens subtilisin to stabilize the same main chain turn. Alterations at this residue should alter the 101-103 main chain direction. Mutations at Glu103 are also expected to affect the 101-103 main chain direction.

The side chain of Gly102 interacts with the substrate P-3 amino acid. Side chains of substituted amino acids thus are expected to significantly affect specificity for the P-3 substrate amino acids.

All the amino acids within the 127-129 segment are considered important to substrate specificity. Gly127 is positioned such that its side chain interacts with the S-1 and S-3 subsites. Altering this residue thus should alter the specificity for P-1 and P-3 residues of the substrate.

The side chain of Gly128 comprises a part of both the S-2 and S-4 subsites. Altered specificity for P-2 and P-4 therefore would be expected upon mutation. Moreover, such mutation may convert subtilisin into an amino peptidase for the same reasons substitutions of Leu126 would be expected to produce that result.

The Pro129 residue is likely to restrict the conformational freedom of the sequence 126-133, residues which may play a major role in determining P-1 specificity. Replacing Pro may introduce more flexibility thereby broadening the range of binding capabilities of such mutants.

The side chain of Lys213 is located within the S-3 subsite. All of the amino acids within the 213-215 segment are also considered to be important to substrate specificity. Accordingly, altered P-3 substrate specificity is expected upon mutation of this residue.

The Tyr214 residue does not interact with substrate but is positioned such that it could affect the conformation of the hair pin loop 204-217.

Finally, mutation of the Gly215 residue should affect the S-3' subsite, and thereby alter P-3' specificity.

In addition to the above substitutions of amino acids, the insertion or deletion of one or more amino acids within the external loop comprising residues 152-172 may also affect specificity. This is because these residues may play a role in the "secondary contact region" described in the model of streptomyces subtilisin inhibitor complexed with subtilisin. Hirono, et al. (1984) J. Mol. Biol. 178, 389-413. Thermitase K has a deletion in this region, which eliminates several of these "secondary contact" residues. In particular, deletion of residues 161 through 164 is expected to produce a mutant subtilisin having modified substrate specificity. In addition, a rearrangement in this area induced by the deletion should alter the position of many residues involved in substrate binding, predominantly at P-1. This, in turn, should affect overall activity against proteinaceous substrates

The effect of deletion of residues 161 through 164 has been shown by comparing the activity of the wild type (WT) enzyme with a mutant enzyme containing this deletion as well as multiple substitutions (i.e., S153/S156/A158/G159/S160/Δ161-164/I165/S166/A169/R170). This produced the following results:

TABLE V

	kcat	Km	kcat/Km
WT	50	1.4×10^{-4}	3.6×10^{5}
Deletion mutant	8	5.0×10^{-6}	1.6×10^{6}

The WT has a kcat 6 times greater than the deletion mutant but substrate binding is 28 fold tighter by the deletion mutant. The overall efficiency of the deletion mutant is thus 4.4 times higher than the WT enzyme.

All of these above identified residues which have yet to be substituted, deleted or inserted into are presented in Table VI.

TABLE VI

Substitution/Insertion/Deletion	
Residues	
His67	Ala152
Leu126	Ala153
Leu135	Gly154
Gly97	Asn155
Asp99	Gly156
Ser101	Gly157
Gly102	Gly160
Glu103	Thr158
Leu126	Ser159
Gly127	Ser161
Gly128	Ser162
Pro129	Ser163
Tyr214	Thr164
Gly215	Val165
Gly166	Gly169
Tyr167	Lys170
Pro168	Tyr171
	Pro172

The following disclosure is intended to serve as a representation of embodiments herein, and should not be construed as limiting the scope of this application. These specific examples disclose the construction of certain of the above identified mutants. The construction of the other mutants, however, is apparent from the disclosure herein and that presented in EPO Publication No. 0130756.

All literature citations are expressly incorporated by reference.

EXAMPLE 1

Identification of Peracid Oxidizable Residues of Subtilisin Q222 and L222

As shown in Figures 6A and 6B, organic peracid oxidants inactivate the mutant subtilisins Met222L and Met222Q (L222 and Q222). This example describes the identification of peracid oxidizable sites in these mutant subtilisins.

First, the type of amino acid involved in peracid oxidation was determined. Except under drastic conditions (Means, G.E., et al. (1971) Chemical Modifications of Proteins, Holden-Day, S.F., CA, pp. 160-162), organic peracids modify only methionine and tryptophan in subtilisin. Difference spectra of the enzyme over the 250nm to 350nm range were determined during an inactivation titration employing the reagent, diperdodecanoic acid (DPDA) as oxidant. Despite quantitative inactivation of the enzyme, no change in absorbance over this wavelength range was noted as shown in Figures 7A and 7B indicating that tryptophan was not oxidized. Fontana, A., et al. (1980) Methods in Peptide and Protein Sequence Analysis - (C. Birr ed.) Elsevier, New York, p. 309. The absence of tryptophan modification implied oxidation of one or more of the remaining methionines of B. amyloliquefaciens subtilisin. See Figure 1.

To confirm this result the recombinant subtilisin Met222F was cleaved with cyanogen bromide (CNBr) both before and after oxidation by DPDA. The peptides produced by CNBr cleavage were analyzed on high resolution SDS-pyridine peptide gels (SPG).

Subtilisin Met222F (F222) was oxidized in the following manner. Purified F222 was resuspended in 0.1 M sodium borate pH 9.5 at 10 mg/ml and was added to a final concentration of 26 diperdodecanoic acid

(DPDA) at 26 mg/ml was added to produce an effective active oxygen concentration of 30 ppm. The sample was incubated for at least 30 minutes at room temperature and then quenched with 0.1 volume of 1 M Tris pH 8.6 buffer to produce a final concentration of 0.1 M Tris pH 8.6). 3mM phenylmethylsulfonyl fluoride (PMSF) was added and 2.5 ml of the sample was applied to a Pharmacia PD10 column equilibrated in 10 mM sodium phosphate pH 6.2, 1 mM PMSF. 3.5 ml of 10 mM sodium phosphate pH6.2, 1mM PMSF was applied and the eluant collected.

F222 and DPDA oxidized F222 were precipitated with 9 volumes of acetone at -20°C. The samples were resuspended at 10 mg/ml in 8M urea in 88% formic acid and allowed to sit for 5 minutes. An equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein) and the samples incubated for 2 hours at room temperature in the dark. Prior to gel electrophoresis, the samples were lyophilized and resuspended at 2-5 mg/ml in sample buffer (1% pyridine, 5% $NaDodSO_4$, 5% glycerol and bromophenol blue) and disassociated at 95°C for 3 minutes.

The samples were electrophoresed on discontinuous polyacrylamide gels (Kyte, J., et al. (1953) Anal. Bioch. 133, 515-522). The gels were stained using the Pharmacia silver staining technique (Sammons, D.W., et al. (1981) Electrophoresis 2 135-141).

The results of this experiment are shown in Figure 8. As can be seen, F222 treated with CNBr only gives nine resolved bands on SPG. However, when F222 is also treated with DPDA prior to cleavage, bands X, 7 and 9 disappear whereas bands 5 and 6 are greatly increased in intensity.

In order to determine which of the methionines were effected, each of the CNBr peptides was isolated by reversed phase HPLC and further characterized. The buffer system in both Solvent A (aqueous) and Solvent B (organic) for all HPLC separations was 0.05% triethylamime/trifloroacetic acid (TEA-TFA). In all cases unless noted, solvent A consisted of 0.05% TEA-TFA in H_2O, solvent B was 0.05% TEA-TFA in 1-propanol, and the flow rate was 0.5 ml/minute.

For HPLC analysis, two injections of 1 mg enzyme digest were used. Three samples were acetone precipitated, washed and dried. The dried 1 mg samples were resuspended at 10 mg/ml in 8M urea, 88% formic acid; an equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein). After incubation for 2 hours in the dark at room temperature, the samples were desalted on a 0.8 cm X 7 cm column of Tris Acryl GF05 coarse resin (IBF, Paris, France) equilibrated with 40% solvent B, 60% solvent A. 200 ul samples were applied at a flow rate of 1 ml a minute and 1.0-1.2 ml collected by monitoring the absorbance at 280nm. Prior to injection on the HPLC, each desalted sample was diluted with 3 volumes of solvent A. The samples were injected at 1.0 ml/min (2 minutes) and the flow then adjusted to 0.5 ml/min (100% A). After 2 minutes, a linear gradient to 60% B at 1.0% B/min was initiated. From each 1 mg run, the pooled peaks were sampled (50ul) and analyzed by gel electrophoresis as described above.

Each polypeptide isolated by reversed phase HPLC was further analyzed for homogeneity by SPG. The position of each peptide on the known gene sequence (Wells, J.A., et al. (1983) Nucleic Acids Res. 11 7911-7924) was obtained through a combination of amino acid compositional analysis and, where needed, amino terminal sequencing.

Prior to such analysis the following peptides were to rechromatographed.

1. CNBr peptides from F222 not treated with DPDA:

Peptide 5 was subjected to two additional reversed phase separations. The 10 cm C4 column was equilibrated to 80%A/ 20%B and the pooled sample applied and washed for 2 minutes. Next an 0.5% ml B/min gradient was initiated. Fractions from this separation were again rerun, this time on the 25 cm C4 column, and employing 0.05% TEA-TFA in acetonitrile/1-propanol (1:1) for solvent B. The gradient was identical to the one just described.

Peptide "X" was subjected to one additional separation after the initial chromatography. The sample was applied and washed for 2 minutes at 0.5ml/min (100%A), and a 0.5% ml B/min gradient was initiated.

Peptides 7 and 9 were rechromatographed in a similar manner to the first rerun of peptide 5.

Peptide 8 was purified to homogeneity after the initial separation.

2. CNBr Peptides from DPDA Oxidized F222:

Peptides 5 and 6 from a CNBr digest of the oxidized F222 were purified in the same manner as peptide 5 from the untreated enzyme.

Amino acid compositional analysis was obtained as follows. Samples (-1nM each amino acid) were dried, hydrolyzed in vacuo with 100 ul 6N HCl at 106°C for 24 hours and then dried in a Speed Vac. The samples were analyzed on a Beckmann 6300 AA analyzer employing ninhydrin detection.

Amino terminal sequence data was obtained as previously described (Rodriguez, H., et al. (1984) Anal. Biochem. 134, 538-547).

The results are shown in Table VII and Figure 9.

TABLE VII

Amino and COOH terminii of CNBr fragments Terminus and Method		
Fragment	amino, method	COOH, method
X	1, sequence	50, composition
9	51, sequence	119, composition
7	125, sequence	199, composition
8	200, sequence	275, composition
5ox	1, sequence	119, composition
6ox	120, composition	199, composition

Peptides 5ox and 6ox refer to peptides 5 and 6 isolated from CNBr digests of the oxidized protein where their respective levels are enhanced.

From the data in Table VII and the comparison of SPG tracks for the oxidized and native protein digests in Figure 8, it is apparent that (1) Met50 is oxidized leading to the loss of peptides X and 9 and the appearance of 5; and (2) Met124 is also oxidized leading to the loss of peptide 7 and the accumulation of peptide 6. Thus oxidation of B. amyloliquifaciens subtilisin with the peracid, diperdocecanoic acid leads to the specific oxidation of methionine at residues 50 and 124.

EXAMPLE 2

Substitution at Met50 and Met124 in Subtilisin Met222Q

The choice of amino acid for substitution at Met50 was based on the available sequence data for subtilisins from B. licheniformis (Smith, E.C., et al. (1968) J. Biol. Chem. 243, 2184-2191), B.DY (Nedkov, P., et al. (1983) Hoppe Sayler's Z. Physiol. Chem. 364 1537-1540), B. amylosacchariticus (Markland, F.S., et al. (1967) J. Biol. Chem. 242 5198-5211) and B. subtilis (Stahl, M.L., et al. (1984) J. Bacteriol. 158, 411-418). In all cases, position 50 is a phenylalanine. See Figure 5. Therefore, Phe50 was chosen for construction.

At position 124, all known subtilisins possess a methionine. See Figure 5. Molecular modelling of the x-ray derived protein structure was therefore rehired to determine the most probable candidates for substitution. From all 19 candidates, isoleucine and leucine were chosen as the best residues to employ. In order to test whether or not modification at one site but not both was sufficient to increase oxidative stability, all possible combinations were built on the Q222 backbone (F50/Q222, I124/Q222, F50/I124/Q222).

A. Construction of Mutations Between Codons 45 and 50

All manipulations for cassette mutagenesis were carried out on pS4.5 using methods disclosed in EPO Publication No. 0130756 and Wells, J.A., et al, (1985) Gene 34, 315-323. The pΔ50 in Fig. 10, line 4, mutations was produced using the mutagenesis primer shown in Fig. 10, line 6, and employed an approach designated as restriction-purification which is described below. Briefly, a M13 template containing the subtilisin gene, M13mp11-SUBT was used for heteroduplex synthesis (Adelman, et al (1983), DNA 2, 183-193). Following transfection of JM101 (ATCC 33876), the 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was subcloned from M13mp11 SUBT rf into a recipient vector fragment of pBS42 the construction of which is described in EPO Publication No. 0130756. To enrich for the mutant sequence (pΔ50, line 4), the resulting plasmid pool was digested with KpnI, and linear molecules were purified by polyacrylamide gel electrophoresis. Linear molecules were ligated back to a circular form, and transformed into E. coli MM294 cells (ATCC 31446). Isolated plasmids were screened by restriction analysis for the KpnI, site. KpnI$^+$ plasmids were sequenced and confirmed the pΔ50 sequence. Asterisks in Figure 11 indicate the bases that are mutated from the wid type sequence (line 4). pΔ50 (line 4) was cut with StuI and EcoRI and the 0.5 Kb fragment containing the 5' half of the subtilisin gene was purified (fragment 1). pΔ50 (line 4) was digested with KpnI and EcoRI and the 4.0 Kb fragment containing the 3' half of the subtilisin gene and vector sequences was purified (fragment 2). Fragments 1 and 2 (line 5), and duplex DNA

cassettes coding for mutations desired (shaded sequence, line 6) were mixed in a molar ratio of 1:1:10, respectively. For the particular construction of this example the DNA cassette contained the triplet TTT for codon 50 which encodes Phe. This plasmid was designated pF50. The mutant subtilisin was designated F50.

B. Construction of Mutation Between Codons 122 and 127

The procedure of Example 2A was followed in substantial detail except that the mutagenesis primer of Figure 11, line 7 was used and restriction-purification for the EcoRV site in pΔ124 was used. In addition, the DNA cassette (shaded sequence, Figure 11, line 6) contained the triplet ATT for codon 124 which encodes Ile and CTT for Leu. Those plasmids which contained the substitution of Ile for Met124were designeated pI124. The mutant subtilisin was designated I124.

C. Construction of Various F50/I124/Q222 Multiple Mutants

The triple mutant, F50/I124/Q222, was constructed from a three-way ligation in which each fragment contained one of the three mutations. The single mutant Q222 (pQ222) was prepared by cassette mutagenesis as described in EPO Publication No. 0130756. The F50 mutation was contained on a 2.2kb AvaII to PvuII fragment from pF50; the I124 mutation was contained on a 260 bp PvuII to AvaII fragment from pI124; and the Q222 mutation was contained on 2.7 kb AvaII to AvaII fragment from pQ222. The three fragments were ligated together and transformed into E. coli MM294 cells. Restriction analysis of plasmids from isolated transformants confirmed the construction. To analyze the final construction it was convenient that the AvaII site at position 798 in the wild-type subtilisin gene was eliminated by the I124 construction.

The F50/Q222 and I124/Q222 mutants were constructed in a similar manner except that the appropriate fragment from pS4.5 was used for the final construction.

D. Oxidative Stability of Q222 Mutants

The above mutants were analyzed for stability to peracid oxidation. As shown in Fig. 12, upon incubation with diperdodecanoic acid (protein 2mg/mL, oxidant 75ppm[0]), both the I124/Q222 and the F50/I124/Q222 are completely stable whereas the F50/Q222 and the Q222 are inactivated. This indicates that conversion of Met124 to I124 in subtilisin Q222 is sufficient to confer resistance to organic peracid oxidants.

EXAMPLE 3

Subtilisin Mutants Having Altered Substrate Specificity-Hydrophobic Substitutions at Residues 166

Subtilisin contains an extended binding cleft which is hydrophobic in character. A conserved glycine at residue 166 was replaced with twelve non-ionic amino acids which can project their side-chains into the S-1 subsite. These mutants were constructed to determine the effect of changes in size and hydrophobicity on the binding of various substrates.

A. Kinetics for Hydrolysis of Substrates Having Altered P-1 Amino Acids by Subtilisin from B. Amyloliquefaciens

Wild-type subtilisin was purified from B. subtilis culture supernatants expressing the B. amyloliquefaciens subtilisin gene (Wells, J.A., et al. (1983) Nucleic Acids Res. 11, 7911-7925) as previously described (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Details of the synthesis of tetrapeptide substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide (where X is the P1 amino acid) are described by DelMar, E.G., et al. (1979) Anal. Biochem. 99, 316-320. Kinetic parameters, Km(M) and kcat-(s^{-1}) were measured using a modified progress curve analysis (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Briefly, plots of rate versus product concentration were fit to the differential form of the rate equation using a non-linear regression algorithm. Errors in kcat and Km for all values reported are less than five percent. The various substrates in Table VIII are ranged in order of decreasing hydrophobicity. Nozaki, Y. (1971), J. Biol. Chem. 246, 2211-2217; Tanford C. (1978) Science 200, 1012).

TABLE VIII

P1 substrate Amino Acid	kcat(S^{-1})	1/Km(M^{-1})	kcat/Km ($s^{-1}M^{-1}$)
Phe	50	7,100	360,000
Tyr	28	40,000	1,100,000
Leu	24	3,100	75,000
Met	13	9,400	120,000
His	7.9	1,600	13,000
Ala	1.9	5,500	11,000
Gly	0.003	8,300	21
Gln	3.2	2,200	7,100
Ser	2.8	1,500	4,200
Glu	0.54	32	16

The ratio of kcat/Km (also referred to as catalytic efficienty) is the apparent second order rate constant for the conversion of free enzyme plus substrate (E+S) to enzyme plus products (E+P) (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287). The log (kcat/Km) is proportional to transition state binding energy, ΔG_T^* . A plot of the log kcat/Km versus the hydrophobicity of the P1 side-chain (Figure 14) shows a strong correlation (r = 0.98), with the exception of the glycine substrate which shows evidence for non-productive binding. These data show that relative differences between transition-state binding energies can be accounted for by differences in P-1 side-chain hydrophobicity. When the transition-state binding energies are calculated for these substrates and plotted versus their respective side-chain hydrophobicities, the line slope is 1.2 (not shown). A slope greater than unity, as is also the case for chymotrypsin (Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Harper, J.W., et al. (1984) Biochemistry, 23, 2995-3002), suggests that the P1 binding cleft is more hydrophobic than ethanol or dioxane solvents that were used to empirically determine the hydrophobicity of amino acids (Nozaki, Y., et al. J. Biol. Chem. (1971) 246, 2211-2217; Tanford, C. (1978) Science 200, 1012).

For amide hydrolysis by subtilisin, kcat can be interpreted as the acylation rate constant and Km as the dissociation constant, for the Michaelis complex (E•S), Ks. Gutfreund, H., et al (1956) Biochem. J. 63, 656. The fact that the log kcat, as well as log 1/Km, correlates with substrate hydrophobicity is consistent with proposals (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303) that during the acylation step the P-1 side-chain moves deeper into the hydrophobic cleft as the substrate advances from the Michaelis complex (E•S) to the tetrahedral transition-state complex (E•S*). However, these data can also be interpreted as the hydrophobicity of the P1 side-chain effecting the orientation, and thus the susceptibility of the scissile peptide bond to nucleophilic attack by the hydroxyl group of the catalytic Ser221.

The dependence of kcat/Km on P-1 side chain hydrophobicity suggested that the kcat/Km for hydrophobic substrates may be increased by increasing the hydrophobicity of the S-1 binding subsite. To test this hypothesis, hydrophobic amino acid substitutions of Gly166 were produced.

Since hydrophobicity of aliphatic side-chains is directly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927), increasing the hydrophobicity in the S-1 subsite may also sterically hinder binding of larger substrates. Because of difficulties in predicting the relative importance of these two opposing effects, we elected to generate twelve non-charged mutations at position 166 to determine the resulting specificities against non-charged substrates of varied size and hydrophobicity.

B. Cassette Mutagenesis of the P1 Binding Cleft

The preparation of mutant subtilisims containing the substitution of the hydrophobic amino acids Ala, Val and Phe into residue 166 has been described in EPO Publication No. 0130756. The same method was used to produce the remaining hydrophobic mutants at residue 166. In applying this method, two unique and silent restriction sites were introduced in the subtilisin genes to closely flank the target codon 166. As can be seen in Figure 13, the wild type sequence (line 1) was altered by site-directed mutagenesis in M13 using the indicated 37mer mutagenesis primer, to introduce a 13 bp delection (dashedline) and unique SacI and XmaI sites (underlined sequences) that closely flank codon 166. The subtilisin gene fragment was subcloned back into the E. coli - B. subtilis shuttle plasmid, pBS42, giving the plasmid pΔ166 (Figure 13,

line 2). pΔ166 was cut open with SacI and XmaI, and gapped linear molecules were purified (Figure 13, line 3). Pools of synthetic oligonucleotides containing the mutation of interest were annealed to give duplex DNA cassettes that were ligated into gapped pΔ166 (underlined and overlined sequences in Figure 13, line 4). This construction restored the coding sequence except over position 166(NNN; line 4). Mutant sequences were confirmed by dideoxy sequencing. Asterisks denote sequence changes from the wild type sequence. Plasmids containing each mutant B. amyloliquefaciens subtilisin gene were expressed at roughly equivalent levels in a protease deficient strain of B. subtilis, BG2036 as previously described. EPO Publication No. 0130756; Yang, M., et al. (1984) J. Bacteriol. 160, 15-21; Estell, D.A., et al (1985) J. Biol. Chem. 260, 6518-6521.

C. Narrowing Substrate Specificity by Steric Hindrance

To probe the change in substrate specificity caused by steric alterations in the S-1 subsite, position 166 mutants were kinetically analyzed versus P1 substrates of increasing size (i.e., Ala, Met, Phe and Tyr). Ratios of kcat/Km are presented in log form in Figure 15 to allow direct comparisons of transition-state binding energies between various enzyme-substrate pairs.

According to transition state theory, the free enery difference between the free enzyme plus substrate (E + S) and the transition state complex (E•S*) can be calculated from equation (1),

$$(1) \quad {}^{\Delta}G_{T}^{\neq} = -RT \ln kcat/Km + RT \ln kT/h$$

in which kcat is the turnover number, Km is the Michaelis constant, R is the gas constant, T is the temperature, k is Boltzmann's constant, and h is Planck's constant. Specificity differences are ezpressed quantitatively as differences between transition state binding energies (i.e., $\Delta\Delta G_T^{\neq}$), and can be calculated from equation (2).

$$(2) \quad {}^{\Delta\Delta}G_{T}^{\neq} = -RT \ln (kcat/Km)_A/(kcat/Km)_B$$

A and B represent either two different substrates assayed againt the same enzyme, or two mutant enzymes assayed against the same substrate.

As can be seen from Figure 15A, as the size of the side-chain at position 166 increases the substrate preference shifts from large to small P-1 side-chains. Enlarging the side-chain at position 166 causes kcat/Km to decrease in proportion to the size of the P-1 substrate side-chain (e.g., from Gly166 (wild-type) through W166, the kcat/Km for the Tyr substrate is decreased most followed in order by the Phe, Met and Ala P-1 substrates).

Specific steric changes in the position 166 side-chain, such as he presence of a β-hydroxyl group, β- or γ-aliphatic branching, cause large decreases in kcat/Km for larger P1 substrates. Introducing a β-hydroxyl group in going from A166 (Figure 15A) to S166 (Figure 15B), causes an 8 fold and 4 fold reduction in kcat/Km for Phe and Tyr substrates, respectively, while the values for Ala and Met substrates are unchanged. Producing a β-branched structure, in going from S166 to T166, results in a drop of 14 and 4 fold in kcat/Km for Phe and Tyr, respectively. These differences are slightly magnified for V166 which is slightly larger and isosteric with T166. Enlarging the β-branched substituents from V166 to I166 causes a lowering of kcat/Km between two and six fold toward Met, Phe and Tyr substrates. Inserting a γ-branched structure, by replacing M166 (Figure 15A) with L166 (Figure 15B), produces a 5 fold and 18 fold decrease in kcat/Km for Phe and Tyr substrates, respectively. Aliphatic γ-branched appears to induce less steric hindrance toward the Phe P-1 substrate than β-branching, as evidenced by the 100 fold decrease in kcat/Km for the Phe substrate in going from L166 to I166.

Reductions in kcat/Km resulting from increases in side chain size in the S-1 subsite, or specific structural features such as β- and γ-branching, are quantitatively illustrated in Figure 16. The kcat/Km values for the position 166 mutants determined for the Ala, Met, Phe, and Tyr P-1 substrates (top panel through bottom panel, respectively), are plotted versus the position 166 side-chain volumes (Chothia, C. (1984) Ann. Rev. Biochem. 53, 537-572). Catalytic efficiency for the Ala substrate reaches a maximum for

I166, and for the Met substrate it reaches a maximum between V166 and L166. The Phe substrate shows a broad kcat/Km peak but is optimal with A166. Here, the β-branched position 166 substitutions form a line that is parallel to, but roughly 50 fold lower in kcat/Km than side-chains of similar size [i.e., C166 versus T166, L166 versus I166). The Tyr substrate is most efficiently utilized by wild type enzyme (Gly166), and there is a steady decrease as one proceeds to large position 166 side-chains. The β-branched and γ-branched substitutions form a parallel line below the other non-charged substitutions of similar molecular volume.

The optimal substitution at position 166 decreases in volume with increasing volume of the P1 substrate [i.e., I166/Ala substrate, L166/Met substrate, A166/Phe substrate, Gly166/Tyr substrate]. The combined volumes for these optimal pairs may approximate the volume for productive binding in the S-1 subsite. For the optimal pairs, Gly166/Tyr substrate, A166/Phe substrate, L166/Met substrate, V166/Met substrate, and I166/Ala substrate, the combined volumes are 266,295,313,339 and 261 A^3, respectively. Subtracting the volume of the peptide backbone from each pair (i.e., two times the volume of glycine), an average side-chain volume of 160±32A^3 for productive binding can be calculated.

The effect of volume, in excess to the productive binding volume, on the drop in transition-state binding energy can be estimated from the Tyr substrate curve (bottom panel, Figure 16), because these data, and modeling studies (Figure 2), suggest that any substitution beyond glycine causes steric repulsion. A best-fit line drawn to all the data ($r = 0.87$) gives a slope indicating a loss of roughly 3 kcal/mol in transition state binding energy per 100A^3 of excess volume. (100A^3 is approximately the size of a leucyl side-chain.)

D. Enhanced Catalytic Efficiency Correlates with Increasing Hydrophobicity of the Position 166 Substitution

Substantial increases in kcat/Km occur with enlargement of the position 166 side-chain, except for the Tyr P-1 substrate (Figure 16). For example, kcat/Km increases in progressing from Gly166 to I166 for the Ala substrate (net of ten-fold), from Gly166 to L166 for the Met substrate (net of ten-fold) and from Gly166 to A166 for the Phe substrate (net of two-fold). The increases in kcat/Km cannot be entirely explained by the attractive terms in the van der Waals potential energy function because of their strong distance dependence ($1/r^6$) and because of the weak nature of these attractive forces (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Levitt, M. (1976) J. Mol. Biol. 104, 59-107). For example, Levitt (Levitt, M. (1976) J. Mol. Biol. 104, 59-107) has calculated that the van der Waals attraction between two methionyl residues would produce a maximal interaction energy of roughly -0.2 kcal/mol. This energy would translate to only 1.4 fold increase in kcat/Km.

The increases of catalytic efficiency caused by side-chain substitutions at position 166 are better accounted for by increases in the hydrophobicity of the S-1 subsite. The increase kcat/Km observed for the Ala and Met substrates with increasing position 166 side-chain size would be expected, because hydrophobicity is roughly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927).

Another example that can be interpreted as a hydrophobic effect is seen when comparing kcat/Km for isosteric substitutions that differ in hydrophobicity such as S166 and C166 (Figure 16). Cysteine is considerably more hydrophobic than serine (-1.0 versus +0.3 kcal/mol) (Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012). The difference in hydrophobicity correlates with the observation that C166 becomes more efficient relative to Ser166 as the hydrophobicity of the substrates increases (i.e., Ala < Met < Tye < Phe). Steric hindrance cannot explain these differences because serine is considerably smaller than cysteine (99 versus 118A^3). Paul, I.C., Chemistry of the -SH Group (ed. S. Patai, Wiley Interscience, New York, 1974) pp. 111-149.

E. Production of an Elastase-Like Specificity in Subtilisin

The I166 mutation illustrates particularly well that large changes in specificity can be produced by altering the structure and hydrophobicity of the S-1 subsite by a single mutation (Figure 17). Progressing through the small hydrophobic substrates, a maximal specificity improvement over wild type occurs for the Val substrate (16 fold in kcat/Km). As the substrate side chain size increases, these enhancements shrink to near unity (i.e., Leu and His substrates). The I166 enzyme becomes poorer against larger aromatic substrates of increasing size (e.g., I166 is over 1,000 fold worse against the Tyr substrate than is Gly166). We interpret the increase in catalytic efficiency toward the small hydrophobic substrates for I166 compared to Gly166 to the greater hydrophobicity of isoluecine (i.e., -1.8 kcal/mol versus 0). Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012. The decrease in catalytic efficiency

toward the very large substrates for I166 versus Gly166 is attributed to steric repulsion.

The specificity differences between Gly166 and I166 are similar to the specificity differences between chymotrypsin and the evolutionary relative, elastase (Harper, J.W., et al (1984) Biochemistry 23, 2995-3002). In elastase, the bulky amino acids, Thr and Val, block access to the P-1 binding site for large hydrophobic substrates that are preferred by chymotrypsin. In addition, the catalytic efficiencies toward small hydrophobic substrates are greater for elastase than for chymotrypsin as we obeseve for I166 versus Gly166 in subtilisin.

EXAMPLE 4

Substitution of Ionic Amino Acids for Gly166

The construction of subtilisin mutants containing the substitution of the ionic amino acids Asp, Asn, Gln, Lys and Ang are disclosed in EPO Publication No. 0130756. The present example describes the construction of the mutant subtilisin containing Glu at position 166 (E166) and presents substrate specificity data on these mutants. Further data on position 166 and 156 single and double mutants is presented infra.

pΔ166, described in Example 3, was digested with SacI and XmaI. The double strand DNA cassette (underlined and overlined) of line 4 in Figure 13 contained the triplet GAA for the codon 166 to encode the replacement of Glu for Gly166. This mutant plasmid designated pQ166 was propagated in BG2036 as described. This mutant subtilisin, together with the other mutants containing ionic substituent amino acids at residue 166, were isolated as described and further analyzed for variations in substrate specificity.

Each of these mutants was analyzed with the tetrapeptide substrates, succinyl-L-AlaL-AlaProL-X-p-nitroanilide, where X was Phe, Ala and Glu.

The results of this analysis are shown in Table IX.

TABLE IX

Position 166	P-1 Substrate (kcat/Km x 10^{-4})		
	Phe	Ala	Glu
Gly (wild type)	36.0	1.4	0.002
Asp (D)	0.5	0.4	<0.001
Glu (E)	3.5	0.4	<0.001
Asn (N)	18.0	1.2	0.004
Gln (Q)	57.0	2.6	0.002
Lys (K)	52.0	2.8	1.2
Arg (R)	42.0	5.0	0.08

These results indicate that charged amino acid substitutions at Gly166 have improved catalytic efficiencies (kcat/Km) for oppositely charged P-1 substrates (as much as 500 fold) and poorer catalytic efficiency for like charged P-1 substrates.

EXAMPLE 5

Substitution of Glycine at Position 169

The substitution of Gly169 in B. amyloliquefaciens subtilisin with Ala and Ser is described in EPO Publication No. 0130756. The same method was used to make the remaining 17 mutants containing all other substituent amino acids for position 169.

The construction protocol is summarized in Figure 18. The overscored and underscored double stranded DNA cassettes used contained the following triplet encoding the substitution of the indicated amino acid at residue 169.

GCT	A	ATG	M
TGT	C	AAC	N
GAT	D	CCT	P
GAA	E	CAA	Q
TTC	F	AGA	R
GGC	G	AGC	S
CAC	H	ACA	T
ATC	I	GTT	V
AAA	K	TGG	W
CTT	L	TAC	Y

Each of the plasmids containing a substituted Gly169 was designated pX169, where X represents the substituent amino acid. The mutant subtilisins were simialrly designated.

Two of the above mutant subtilisins, A169 and S169, were analyzed for substrate specificity against synthetic substrates containing Phe, Leu, Ala and Arg in the P-1 position. The following results are shown in Table X.

TABLE X

Effect of Serine and Alanine Mutations at Position 169 on P-1 Substrate Specificity				
Position 169	P-1 Substrate [kcat/Km x 10^{-4})			
	Phe	Leu	Ala	Arg
Gly (wild type)	40	10	1	0.4
A169	120	20	1	0.9
S169	50	10	1	0.6

These results indicate that substitutions of Ala and Ser at Gly169 have remarkably similar catalytic efficiencies against a range of P-1 substrates compared to their position 166 counterparts. This is probably because position 169 is at the bottom of the P-1 specificity subsite.

EXAMPLE 6

Substitution at Position 104

Tyr104 has been substituted with Ala, His, Leu, Met and Ser. The method used was a modification of the site directed mutagenesis method. According to the protocol of Figure 19, a primer (shaded in line 4) introduced a unique HindIII site and a frame shift mutation at codon 104. Restriction-purification for the unique HindIII site facilitated the isolation of the mutant sequence (line 4). Restriction-selection against this HindIII site using pimers in line 5 was used to obtain position 104 mutants.

The following triplets were used in the primers of Figure 19, line 5 for the 104 codon which substituted the following amino acids.

GCT	A	TTC	F
ATG	M	CCT	P
CTT	L	ACA	T
AGC	S	TGG	W
CAC	H	TAC	Y
CAA	Q	GTT	V
GAA	E	AGA	R
GGC	G	AAC	N
ATC	I	GAT	D
AAA	K	TGT	C

The substrates in Table XI were used to analyze the substrate specificity of these mutants. The results obtained fo H104 subtilisin are shown in Table XI.

TABLE XI

Substrate	kcat		Km		Kcat/Km	
	WT	H104	WT	H104	WT	H104
sAAPFpNA	50.0	22.0	$1.4x10^{-4}$	$7.1x10^{-4}$	$3.6x10^{5}$	$3.1x10^{4}$
sAAPApNA	3.2	2.0	$2.3x10^{-4}$	$1.9x10^{-3}$	$1.4x10^{4}$	$1x10^{3}$
sFAPFpNA	26.0	38.0	$1.8x10^{-4}$	$4.1x10^{-4}$	$1.5x10^{5}$	$9.1x10^{4}$
sFAPApNA	0.32	2.4	$7.3x10^{-5}$	$1.5x10^{-4}$	$4.4x10^{3}$	$1.6x10^{4}$

From these data it is clear that the substitution of His for Tyr at position 104 produces an enzyme which is more efficient (higher kcat/Km) when Phe is at the P-4 substrate position than when Ala is at the P-4 substrate position.

EXAMPLE 7

Substitution of Ala152

Ala152 has been substituted by Gly and Ser to determine the effect of such substitutions on substrate specificity.

The wild type DNA sequence was mutated by the V152/P153 primer (Figure 20, line 4) using the above restriction-purification approach for the new KpnI site. Other mutant primers (shaded sequences Figure 20; S152, line 5 and G152, line 6) mutated the new KpnI site away and such mutants were isolated using the restriction-selection procedure as described above for loss of the KpnI site.

The results of these substitutions for the above synthetic substrates containing the P-1 amino acids Phe, Leu and Ala are shown in Table XII.

TABLE XII

Position 152	P-1 Substrate (kcat/Km$x10^{-4}$)		
	Phe	Leu	Ala
Gly (G)	0.2	0.4	<0.04
Ala (wild type)	40.0	10.0	1.0
Ser (S)	1.0	0.5	0.2

These results indicate that, in contrast to positions 166 and 169, replacement of Ala152 with Ser or Gly causes a dramatic reduction in catalytic efficiencies across all substrates tested. This suggests Ala152, at the top of the S-1 subsite, may be the optimal amino acid because Ser end Gly ore homologous Ala substitutes.

EXAMPLE 8

Substitution at Position 156

Mutants containing the substitution of Ser and Gln for Glu156 have been constructed according to the overall method depicted in Figure 21. This method was designed to facilitate the construciton of multiple mutants at position 156 and 166 as will be described hereinafter. However, by regenerating the wild type Gly166, single mutations at Glu156 were obtained.

The plasmid pΔ166 is already depicted in line 2 of Figure 13. The synthetic oligonucleotides at the top right of Figure 21 represent the same DNA cassettes depicted in line 4 of Figure 13. The plasmid p166 in Figure 21 thus represents the mutant plasmids of Examples 3 and 4. In this particular example, p166 contains the wild type Gly166.

Construction of position 156 single mutants were prepared by ligation of the three fragments (1-3) indicated at the bottom of Figure 21. Fragment 3, containing the carboxy-terminal portion of the subtilisin gene including the wild type position 166 codon, was isolated as a 610 bp SacI-BamHI fragment. Fragment 1 contained the vector sequences, as well as the amino-terminal sequences of the subtilisin gene through codon 151. To produce fragment 1, a unique KpnI site at codon 152 was introduced into the wild type subtilisin sequence from pS4.5. Site-directed mutagenesis in M13 employed a primer having the sequence 5'-TA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3' to produce the mutation. Enrichment for the mutant sequence was accomplished by restriction with KpnI, purification and self ligation. The mutant sequence containing the KpnI site was confirmed by direct plasmid sequencing to give pV152. pV152 (~1 µg) was digested with KpnI and treated with 2 units of DNA polymerase I large fragment (Klenow fragment from Boeringer-Mannheim) plus 50 µM deoxynucleotide triphosphates at 37°C for 30 min. This created a blunt end that terminated with codon 151. The DNA was extracted with 1:1 volumes phenol and $CHCl_3$ and DNA in the aqueous phase was precipitated by addition of 0.1 volumes 5M ammonium acetate and two volumes ethanol. After centrifugation and washing the DNA pellet with 70% ethanol, the DNA was lyophilized. DNA was digested with BamHI and the 4.6kb piece (fragment 1) was purified by acrylamide gel electrophoresis followed by electroelution. Fragment 2 was a duplex synthetic DNA cassette which when ligated with fragments 1 and 3 properly restored the coding sequence except at codon 156. The top strand was synthesized to contain a glutamine codon, and the complementary bottom strand coded for serine at 156. Ligation of heterophosphorylated cassettes leads to a large and favorable bias for the phosphorylated over the non-phosphorylated oligonucleotide sequence in the final segrated plasmid product. Therefore, to obtain Q156 the top strand was phosphorylated, and annealed to the non-phosphorylated bottom strand prior to ligation. Similarly, to obtain S156 the bottom strand was phosphorylated and annealed to the non-phosphorylated top strand. Mutant sequences were isolated after ligation and transformation, and were confirmed by restriction analysis and DNA sequencing as before. To express variant subtilisins, plasmids were transformed into a subtilisin-neutral protease deletion mutant of B. subtilis, BG2036, as previously described. Cultures were fermented in shake flasks for 24 h at 37°C in LB media containing 12.5 mg/mL chloramphenicol and subtilisin was purified from culture supernatants as described. Purity of subtilisin was greater than 95% as judged by SDS PAGE.

These mutant plasmids designated pS156 and pQ156 and mutant subtilisins designated S156 and Q156 were analyzed with the above synthetic substrates where P-1 comprised the amino acids Glu, Gln, Met and Lys. The results of this analyses are presented in Example 9.

EXAMPLE 9

Multiple Mutants With Altered Substrate Specificity - Substitution at Positions 156 and 166

Single substitutions of position 166 are described in Examples 3 and 4. Example 8 describes single substitutions at position 156 as well as the protocol of Figure 21 whereby various double mutants comprising the substitution of various amino acids at positions 156 and 166 can be made. This example describes the construction and substrate specificity of subtilisin containing substitutions at position 156 and 166 and summarizes some of the data for single and double mutants at positions 156 and 166 with various substrates.

K166 is a common replacement amino acid in the 156/166 mutants described herein. The replacement of Lys for Gly166 was achieved by using the synthetic DNA cassette at the top right of Figure 21 which contained the triplet AAA for NNN. This produced fragment 2 with Lys substituting for Gly166.

The 156 substituents were Gln and Ser. The Gln and Ser substitutions at Gly156 are contained within fragment 3 (bottom right Figure 21).

The multiple mutants were produced by combining fragments 1, 2 and 3 as described in Example 8. The mutants Q156/K166 and S156/K166 were selectively generated by differential phosphorylation as described. Alternatively, the double 156/166 mutants, c.f. Q156/K166 and S156/K166, were prepared by ligation of the 4.6kb SacI-BamHI fragment from the relevant p156 plasmid containing the 0.6kb SacI-BamHI fragment from the relevant p166 plasmid.

These mutants, the single mutant K166, and the S156 and Q156 mutants of Example 8 were analyzed for substitute specificity against synthetic polypeptides containing Phe or Glu as the P-1 substrate residue. The results are presented in Table XIII.

TABLE XIII

Enzymes Compared[b]	Substrate P-1 Residue	kcat	Km	kcat/Km	kcat/Km (mutant) kcat/Km(wt)
Glu156/Gly166 (WT)	Phe	50.00	$1.4x10^{-4}$	$3.6x10^{5}$	(1)
	Glu	0.54	$3.4x10^{-2}$	$1.6x10^{1}$	(1)
K166	Phe	20.00	$4.0x10^{-5}$	$5.2x10^{5}$	1.4
	Glu	0.70	$5.6x10^{-5}$	$1.2x10^{4}$	750
Q156/K166	Phe	30.00	$1.9x10^{-5}$	$1.6x10^{6}$	4.4
	Glu	1.60	$3.1x10^{-5}$	$5.0x10^{4}$	3100
S156/K166	Phe	30.00	$1.8x10^{-5}$	$1.6x10^{6}$	4.4
	Glu	0.60	$3.9x10^{-5}$	$1.6x10^{4}$	1000
S156	Phe	34.00	$4.7x10^{-5}$	$7.3x10^{5}$	2.0
	Glu	0.40	$1.8x10^{-3}$	$1.1x10^{2}$	6.9
E156	Phe	48.00	$4.5x10^{-5}$	$1.1x10^{6}$	3.1
	Glu	0.90	$3.3x10^{-3}$	$2.7x10^{2}$	17

As can be seen in Table XIV, either of these single mutations improve enzyme performance upon substrates with glutamate at the P-1 enzyme binding site. When these single mutations were combined, the resulting multiple enzyme mutants are better than either parent. These single or multiple mutations also alter the relative pH activity profiles of the enzymes as shown in Figure 23.

To isolate the contribution of electrostatics to substrate specificity from other chemical binding forces, these various single and double mutants were analyzed for their ability to bind and cleave synthetic substrates containing Glu, Gln, Met and Lys as the P-1 substrate amino acid. This permitted comparisons between side-chains that were more sterically similar but differed in charge (e.g., Glu versus Gln, Lys versus Met). Similarly, mutant enzymes were assayed against homologous P-1 substrates that were most sterically similar but differed in charge (Table XIV).

TABLE XIV

Kinetics of Position 156/166 Subtilisins Determined for Different P1 Substrates

Enzyme Position[a]		Net Charge[b]	P-1 Substrate log kcat/Km (log 1/Km)[c]			
			Glu	Gln	Met	Lys
156	166					
Glu	Asp	-2	n.d.	3.02 (2.56)	3.93 (2.74)	4.23 (3.00)
Glu	Glu	-2	n.d.	3.06 (2.91)	3.86 (3.28)	4.48 (3.69)
Glu	Asn	-1	1.62 (2.22)	3.85 (3.14)	4.99 (3.85)	4.15 (2.88)
Glu	Gln	-1	1.20 (2.12)	4.36 (3.64)	5.43 (4.36)	4.10 (3.15)
Gln	Asp	-1	1.30 (1.79)	3.40 (3.08)	4.94 (3.87)	4.41 (3.22)
Ser	Asp	-1	1.23 (2.13)	3.41 (3.09)	4.67 (3.68)	4.24 (3.07)
Glu	Met	-1	1.20 (2.30)	3.89 (3.19)	5.64 (4.83)	4.70 (3.89)
Glu	Ala	-1	n.d.	4.34 (3.55)	5.65 (4.46)	4.90 (3.24)
Glu	Gly(wt)	-1	1.20 (1.47)	3.85 (3.35)	5.07 (3.97)	4.60 (3.13)
Gln	Gly	0	2.42 (2.48)	4.53 (3.81)	5.77 (4.61)	3.76 (2.82)
Ser	Gly	0	2.31 (2.73)	4.09 (3.68)	5.61 (4.55)	3.46 (2.74)
Gln	Asn	0	2.04 (2.72)	4.51 (3.76)	5.79 (4.66)	3.75 (2.74)
Ser	Asn	0	1.91 (2.78)	4.57 (3.82)	5.72 (4.64)	3.68 (2.80)
Glu	Arg	0	2.91 (3.30)	4.26 (3.50)	5.32 (4.22)	3.19 (2.80)
Glu	Lys	0	4.09 (4.25)	4.70 (3.88)	6.15 (4.45)	4.23 (2.93)
Gln	Lys	+1	4.70 (4.50)	4.64 (3.68)	5.97 (4.68)	3.23 (2.75)
Ser	Lys	+1	4.21 (4.40)	4.84 (3.94)	6.16 (4.90)	3.73 (2.84)
Maximum difference: log kcat/Km (log 1/Km)[d]			3.5 (3.0)	1.8 (1.4)	2.3 (2.2)	-1.3 (-1.0)

Footnotes to Table XIV:

(a) B. subtilis, BG 2036, expressing indicated variant subtilisin were fermented and enzymes purified as previously described (Estell, et al. (1985) J. Biol. Chem. 260, 6518-6521). Wild type subtilisin is indicated (wt) containing Glu156 and Gly166.

(b) Net charge in the P-1 binding site is defined as the sum of charges from positions 156 and 166 at pH 8.6.

(c) Values for kcat(s^{-1}) and Km(M) were measured in 0.1M Tris pH 8.6 at 25°C as previously described against P-1 substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide, where X is the indicated P-1 amino acid. Values for log 1/Km are shown inside parentheses. All errors in determination of kcat/Km and 1/Km are below 5%.

(d) Because values for Glu156/Asp166(D166) are too small to determine accurately, the maximum difference taken for GluP-1 substrate is limited to a charge range of +1 to -1 charge change.

n.d. = not determined

The kcat/Km ratios shown are the second order rate constants for the conversion of substrate to product, and represent the catalytic efficiency of the enzyme. These ratios are presented in logarithmic form to scale the data, and because log kcat/Km is proportional to the lowering of transition-state activation energy (ΔG_T). Mutations at position 156 and 166 produce changes in catalytic efficiency toward Glu, Gln, Met and Lys P-1 substrates of 3100, 60, 200 and 20 fold, respectively. Making the P-1 binding-site more positively charged [e.g., compare Gln156/Lys166 (Q156/K166) versus Glu156/Met166 (Glu156/M166)] dramatically increased kcat/Km toward the Glu P-1 substrate (up to 3100 fold), and decreased the catalytic efficiency toward the Lys P-1 substrate (up to 10 fold). In addition, the results show that the catalytic efficiency of wild type enzyme can be greatly improved toward any of the four P-1 substrates by mutagenesis of the P-1 binding site.

The changes in kcat/Km ore caused predominantly by changes in 1/Km. Because 1/Km is approximately equal to 1/Ks, the enzyme-substrate association constant, the mutations primarily cause a change in substrate binding. These mutations produce smaller effects on kcat that run parallel to the effects on 1/Km. The changes in kcat suggest either an alteration in binding in the P-1 binding site in going from the Michaelis-complex E•S) to the transition-state complex (E-S≠) as previously proposed (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), or change in the position of the scissile peptide bond over the catalytic serine in the E•S complex.

Changes in substrate preference that arise from changes in the net charge in the P-1 binding site show trends that are best accounted for by electrostatic effects (Figure 28). As the P-1 binding cleft becomes more positively charged, the average catalytic efficiency increases much more for the Glu P-1 substrate than for its neutral and isosteric P-1 homolog, Gln (Figure 28A). Furthermore, at the positive extreme both substrates have nearly identical catalytic efficiencies.

In contrast, as the P-1 site becomes more positively charged the catalytic efficiency toward the Lys P-1 substrate decreases, and diverges sharply from its neutral and isosteric homolog, Met (Figure 28B). The similar and parallel upward trend seen with increasing positive charge for the Met and Glu P-1 substrates probably results from the fact that all the substrates are succinylated on their amino-terminal end, and thus carry a formal negative charge.

The trends observed in log kcat/Km are dominated by changes in the Km term (Figures 28C and 28D). As the pocket becomes more positively charged, the log 1/Km values converge for Glu and Gln P-1 substrates (Figure 28C), and diverge for Lys and Met P-1 substrates (Figure 28D). Although less

pronounced effects are seen in log kcat, the effects of P-1 charge on log kcat parallel those seen in log 1/Km and become larger as the P-1 pocket becomes more positively charged. This may result from the fact that the transition-state is a tetrahedral anion, and a net positive charge in the enzyme may serve to provide some added stabilization to the transition-state.

The effect of the change in P-1 binding-site charge on substrate preference can be estimated from the differences in slopes between the charged and neutral isosteric P-1 substrates (Figure 28B). The average change in substrate preference (Δlog kcat/Km) between charged and neutral isosteric substrates increases roughly 10-fold as the complementary charge or the enzyme increases (Table XV). When comparing Glu versus Lys, this difference is 100-fold and the change in substrate preference appears predominantly in the Km term.

TABLE XV

Differential Effect on Binding Site Charge on log kcat/Km or (log 1/Km) for P-1 Substrates that Differ in Charge[a]			
Change in P-1 Binding Site Charge[b]	Δlog kcat/Km (Δlog 1/Km)		
	GluGln	MetLys	GluLys
-2 to -1	n.d.	1.2 (1.2)	n.d.
-1 to 0	0.7 (0.6)	1.3 (0.8)	2.1 (1.4)
0 to +1	1.5 (1.3)	0.5 (0.3)	2.0 (1.5)
Avg. change in log kcat/K_m or (log 1/Km) per unit charge change	1.1 (1.0)	1.0 (0.8)	2.1 (1.5)

[a] The difference in the slopes of curves were taken between the P-1 substrates over the charge interval given for log (kcat/Km) (Figure 28A, B) and (log 1/Km) (Figure 28C, D). Values represent the differential effect a charge change has in distinguishing the substrates that are compared.
[b] Charge in P-1 binding site is defined as the sum of charges from positions 156 and 166.

The free energy of electrostatic interactions in the structure and energetics of salt-bridge formation depends on the distance between the charges and the microscopic dielectric of the media. To dissect these structural and microenvironmental effects, the energies involved in specific salt-bridges were evaluated. In addition to the possible salt-bridges shown (Figures 29A and 29B), reasonable salt-bridges can be built between a Lys P-1 substrate and Asp at position 166, and between a Glu P-1 substrate and a Lys at position 166 (not shown). Although only one of these structures is confirmed by X-ray crystalography (Poulos, T.L., et al. (1976) J. Mol. Biol. 257 1097-1103), all models have favorable torsion angles (Sielecki, A.R., et al. (1979) J. Mol. Biol. 134, 781-804), and do not introduce unfavorable van der Waals contacts.

The change in charged P-1 substrate preference brought about by formation of the model salt-bridges above are shown in Table XVI.

TABLE XVI

Effect of Salt Bridge Formation Between Enzyme and Substrate on P1 Substrate Preference[a]

Enzymes Compared[b] 1	Enzymes Compared[b] 2	Enzyme Position Changed	P-1 Substrates Compared	Substrate[d] Preference Δlog (kcat/Km) 1	Substrate[d] Preference Δlog (kcat/Km) 2	Change in Substrate Preference ΔΔlog (kcat/Km) (1-2)
Glu156/Asp166	Gln156/Asp166	156	LysMet	+0.30	-0.53	0.83
Glu156/Asn166	Gln156/Asn166	156	LysMet	-0.84	-2.04	1.20
Glu156/Gly166	Gln156/Gly166	156	LysMet	-0.47	-2.10	1.63
Glu156/Lsy-166	Gln156/Lys166	156	LysMet	-1.92	-2.74	0.82
				Ave ΔΔlog (kcat/Km)		1.10 ± 0.3
Glu156/Asp166	Glu156/Asn166	166	LysMet	+0.30	-0.84	1.14
Glu156/Glu166	Glu156/Glu166	166	LysMet	+0.62	-1.33	1.95
Gln156/Asp166	Gln156/Asn166	166	LysMet	-0.53	-2.04	1.51
Ser156/Asp166	Ser156/Asn166	166	LysMet	-0.43	-2.04	1.61
Glu156/Lys166	Glu156/Met166	166	GluGln	-0.63	-2.69	2/06
				Ave ΔΔlog (kcat/Km)		1.70 ± 0.3

EP 0 251 446 B1

Footnotes to Table XVI:

(a) Molecular modeling shows it is possible to form a salt bridge between the indicated charged P-1 substrate and a complementary charge in the P-1 binding site of the enzyme at the indicated position changed.

(b) Enzymes compared have sterically similar amino acid substitutions that differ in charge at the indicated position.

(c) The P-1 substrates compared are structurally similar but differ in charge. The charged P-1 substrate is complementary to the charge change at the position indicated between enzymes 1 and 2.

(d) Date from Table XIV was used to compute the difference in log (kcat/Km) between the charged and the non-charged P-1 substrate (i.e., the substrate preference). The substrate preference is shown separately for enzyme 1 and 2.

(e) The difference in substrate preference between enzyme 1 (more highly charged) and enzyme 2 (more neutral) represents the rate change accompanying the electrostatic interaction.

The difference between catalytic efficiencies (i.e., Δlog kcat/Km) for the charged and neutral P-1 substrates (e.g., Lys minus Met or Glu minus Gln) give the substrate preference for each enzyme. The change in substrate preference (ΔΔlog kcat/Km) between the charged and more neutral enzyme homologs (e.g., Glu156/Gly166 minus Gln156(Q156)/Gly166) reflects the change in catalytic efficiency that may be attributed solely to electrostatic effects.

These results show that the average change in substrate preference is considerably greater when electrostatic substitutions are produced at position 166 (50-fold in kcat/Km) versus position 156 (12-fold in kcat/Km). From these ΔΔlog kcat/Km values, an average change in transition-state stabilization energy can be calculated of -1.5 and -2.4 kcal/mol for substitutions at positions 156 and 166, respectively. This should represent the stabilization energy contributed from a favorable electrostatic interaction for the binding of free enzyme and substrate to form the transition-state complex.

EXAMPLE 10

Substitutions at Position 217

Tyr217 has been substituted by all other 19 amino acids. Cassette mutagenesis as described in EPO publication No. 0130756 was used according to the protocol of Figure 22. The EcoRV restriction site was used for restriction-purification of pΔ217.

Since this position is involved in substrate binding, mutations here effect kinetic parameters of the enzyme. An example is the substitution of Leu for Tyr at position 217. For the substrate sAAPFpNa, this mutant has a kcat of 277 5' and a Km of 4.7×10^{-4} with a kcat/Km ratio of 6×10^{5}. This represents a 5.5-fold increase in kcat with a 3-fold increase in Km over the wild type enzyme.

In addition, replacement of Tyr217 by Lys, Arg, Phe or Leu results in mutant enzymes which are more stable at pHs of about 9-11 than the WT enzyme. Conversely, replacement of Tyr217 by Asp, Glu, Gly or Pro results in enzymes which are less stable at pHs of about 9-11 than the WT enzyme.

EXAMPLE 11

Multiple Mutants Having Altered Thermal Stability

B. amyloliquefacien subtilisin does not contain any cysteine residues. Thus, any attempt to produce thermal stability by Cys cross-linkage required the substitution of more than one amino acid in subtilisin with Cys. The following subtilisin residues were multiply substituted with cysteine:

Thr22/Ser87

Ser24/Ser87

Mutagenesis of Ser24 to Cys was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

5'-pC-TAC-ACT-GGA-T*G*C-AAT-GTT-AAA-G-3'.

(Asterisks show the location of mismatches and the underlined sequence shows the position of the altered Sau3A site.) The B. amyloliquefaciens subtilisin gene on a 1.5 kb EcoRI-BAMHI fragment from pS4.5 was cloned into M13mp11 and single stranded DNA was isolated. This template (M13mp11SUBT) was double primed with the 5' phosphorylated M13 universal sequencing primer and the mutagenesis primer. Adelman, et al. (1983) DNA 2, 183-193. The heteroduplex was transfected into competent JM101 cells and plaques were probed for the mutant sequence (Zoller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500; Wallace, et al. (1981) Nucleic Acid Res. 9, 3647-3656) using a tetramethylammonium chloride hybridization protocol (Wood, et al. (1985) Proc. Natl. Acad. Sci. USA 82, 1585-1588). The Ser87 to Cys mutation was prepared in a similar fashion using a 5' phosphorylated primer having the sequence

5'-pGGC-GTT-GCG-CCA-*TGC-GCA-TCA-CT-3'.

(The asterisk indicates the position of the mismatch and the underlined sequence shows the position of a new MstI site.) The C24 and C87 mutations were obtained at a frequency of one and two percent, respectively. Mutant sequences were confirmed by dideoxy sequencing in M13.

Mutagenesis of Tyr21/Thr22 to A21/C22 was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

5'-pAC-TCT-CAA-GGC-*G*C*T-T*G*T-GG*C-TCA-AAT-GTT-3'.

(The asterisks show mismatches to the wild type sequence and the underlined sequence shows the position of an altered Sau3A site.) Manipulations for heteroduplex synthesis were identical to those described for C24. Because direct cloning of the heteroduplex DNA fragment can yield increased frequencies of mutagenesis, the EcoRI-BamHI subtilisin fragment was purified and ligated into pBS42. E. coli MM 294 cells were transformed with the ligation mixture and plasmid DNA was purified from isolated transformants. Plasmid DNA was screened for the loss of the Sau3A site at codon 23 that was eliminated by the mutagenesis primer. Two out of 16 plasmid preparations had lost the wild type Sau3A site. The mutant sequence was confirmed by dideoxy sequencing in M13.

Double mutants, C22/C87 and C24/C87, were constructed by ligating fragments sharing a common ClaI site that separated the single parent cystine codons. Specifically, the 500 bp EcoRI-ClaI fragment containing the 5' portion of the subtilisin gene (including codons 22 and 24) was ligated with the 4.7 kb ClaI-EcoRI fragment that contained the 3' portion of the subtilisin gene (including codon 87) plus pBS42 vector sequence. E. coli MM 294 was transformed with ligation mixtures and plasmid DNA was purified from individual transformants. Double-cysteine plasmid constructions were identified by restriction site markers originating from the parent cysteine mutants (i.e., C22 and C24, Sau3A minus; Cys87, MstI plus). Plasmids from E. coli were transformed into B. subtilis BG2036. The thermal stability of these mutants as compared to wild type subtilisin are presented in Figure 30 and Tables XVII and XVIII.

TABLE XVII

Effect of DTT on the Half-Time of Autolytic Inactivation of Wild-Type and Disulfide Mutants of Subtilisin*			
Enzyme	$t_{\frac{1}{2}}$		-DTT/+DTT
	-DDT	+DTT	
	min		
Wild-type	95	85	1.1
C22/C87	44	25	1.8
C24/C87	92	62	1.5

(*) Purified enzymes were either treated or not treated with 25mM DTT and dialyzed with or without 10mM DTT in 2mM $CaCl_2$, 50mM Tris (pH 7.5) for 14 hr. at 4 °C. Enzyme concentrations were adjusted to 80µl aliquots were quenched on ice and assayed for residual activity. Half-times for autolytic inactivation were determined from semi-log plots of $\log_{10}$ (residual activity) versus time. These plots were linear for over 90% of the inactivation.

TABLE XVIII

Effect of Mutations in Subtilisin on the Half-Time of Autolytic Inactivation at 58 °C*	
Enzyme	$t_{\frac{1}{2}}$
	min
Wild-type	120
C22	22
C24	120
C87	104
C22/C87	43
C24/C87	115

(*) Half-times for autolytic inactivation were determined for wild-type and mutant subtilisins as described in the legend to Table III. Unpurified and non-reduced enzymes were used directly from B. subtilis culture supernatants.

The disulfides introduced into subtilisin did not improve the autolytic stability of the mutant enzymes when compared to the wild-type enzyme. However, the disulfide bonds did provide a margin of autolytic stability when compared to their corresponding reduced double-cysteine enzyme. Inspection of a highly refined x-ray structure of wild-type B. amyloliquefaciens subtilisin reveals a hydrogen bond between Thr22 and Ser87. Because cysteine is a poor hydrogen donor or acceptor (Paul, I.C. (1974) in Chemistry of the -SH Group (Patai, S., ed.) pp. 111-149, Wiley Interscience, New York) weakening of 22/87 hydrogen bond may explain why the C22 and C87 single-cysteine mutant proteins are less autolytically stable than either C24 or wild-type (Table XVIII). The fact that C22 is less autolytically stable than C87 may be the result of the Tyr21A mutation (Table XVIII). Indeed, construction and analysis of Tyr21/C22 shows the mutant protein has an autolytic stability closer to that of C87. In summary, the C22 and C87 of single-cysteine mutations destabilize the protein toward autolysis, and disulfide bond formation increases the stability to a level less than or equal to that of wild-type enzyme.

EXAMPLE 12

Multiple Mutants Containing Substitutions at Position 222 and Position 166 or 169

Double mutants 166/222 and 169/222 were prepared by ligating together (1) the 2.3kb AcaII fragment from pS4.5 which contains the 5' portion of the subtilisin gene and vector sequences, (2) the 200bp AvaII fragment which contains the relevant 166 or 169 mutations from the respective 166 or 169 plasmids, and (3) the 2.2kb AvaII fragment which contains the relevant 222 mutation 3' and of the subtilisin genes and vector

sequence from the respective p222 plasmid.

Although mutations at position 222 improve oxidation stability they also tend to increase the Km. An example is shown in Table XIX. In this case the A222 mutation was combined with the K166 mutation to give an enzyme with kcat and Km intermediate between the two parent enzymes.

TABLE XIX

	kcat	Km
WT	50	1.4×10^{-4}
A222	42	9.9×10^{-4}
K166	21	3.7×10^{-5}
K166/A222	29	2.0×10^{-4}
substrate sAAPFpNa		

EXAMPLE 13

Multiple Mutants Containing Substitutions at Positions 50, 156, 166, 217 and Combinations Thereof

The double mutant S156/A169 was prepared by ligation of two fragments, each containing one of the relevant mutations. The plasmid pS156 was cut with XmaI and treated with S1 nuclease to create a blunt end at codon 167. After removal of the nuclease by phenol/chloroform extraction and ethanol precipitation, the DNA was digested with BamHI and the approximately 4kb fragment containing the vector plus the 5' portion of the subtilisin gene through codon 167 was purified.

The pA169 plasmid was digested with KpnI and treated with DNA polymerase Klenow fragment plus 50 µM dNTPs to create a blunt end codon at codon 168. The Klenow was removed by phenol/chloroform extraction and ethanol precipitation. The DNA was digested with BamHI and the 590bp fragment including codon 168 through the carboxy terminus of the subtilisin gene was isolated. The two fragments were then ligated to give S156/A169.

Triple and quadruple mutants were prepared by ligating together (1) the 220bp PvuII/HaeII fragment containing the relevant 156, 166 and/or 169 mutations from the respective p156, p166 and/or p169 double of single mutant plasmid, (2) the 550bp HaeII/BamHI fragment containing the relevant 217 mutant from the respective p217 plasmid, and (3) the 3.9kb PvuII/BamHI fragment containing the F50 mutation and vector sequences.

The multiple mutant F50/S156/A169/L217, as well as B. amyloliquefaciens subtilisin, B. lichenformis subtilisin and the single mutant L217 were analyzed with the above synthetic polypeptides where the P-1 amino acid in the substrate was Lys, His, Ala, Gln, Tyr, Phe, Met and Leu. These results are shown in Figures 26 and 27.

These results show that the F50/S156/A169/L217 mutant has substrate specificity similar to that of the B. licheniformis enzyme and differs dramatically from the wild type enzyme. Although only data for the L217 mutant are shown, none of the single mutants (e.g., F50, S156 or A169) showed this effect. Although B. licheniformis differs in 88 residue positions from B. amyloliquefaciens, the combination of only these four mutations accounts for most of the differences in substrate specificity between the two enzymes.

EXAMPLE 14

Subtilisin Mutants Having Altered Alkaline Stability

A random mutagenesis technique was used to generate single and multiple mutations within the B. amyloliquefaciens subtilisin gene. Such mutants were screened for altered alkaline stability. Clones having increased (positive) alkaline stability and decreased (negative) alkaline stability were isolated and sequenced to identify the mutations within the subtilisin gene. Among the positive clones, the mutants V107 and R213 were identified. These single mutants were subsequently combined to produce the mutant V107/R213.

One of the negative clones (V50) from the random mutagenesis experiments resulted in a marked decrease in alkaline stability. Another mutant (P50) was analyzed for alkaline stability to determine the effect

of a different substitution at position 50. The F50 mutant was found to have a greater alkaline stability than wild type subtilisin and when combined with the double mutant V107/R213 resulted in a mutant having an alkaline stability which reflected the aggregate of the alkaline stabilities for each of the individual mutants.

The single mutant R204 and double mutant C204/R213 were identified by alkaline screening after random cassette mutagenesis over the region from position 197 to 228. The C204/R213 mutant was thereafter modified to produce mutants containing the individual mutations C204 and R213 to determine the contribution of each of the individual mutations. Cassette mutagenesis using pooled oligonucleotides to substitute all amino acids at position 204, was utilized to determine which substitution at position 204 would maximize the increase in alkaline stability. The mutation from Lys213 to Arg was maintained constant for each of these substitutions at position 204.

A. Construction of pB0180, an E. coli-B. subtilis Shuttle Plasmid

The 2.9 kb EcoRI-BamHI fragment from pBR327 (Covarrubias, L., et al. (1981) Gene 13, 25-35) was ligated to the 3.7kb EcoRI-BamHI fragment of pBD64 (Gryczan, T., et al. (1980) J. Bacteriol., 141, 246-253) to give the recombinant plasmid pB0153. The unique EcoRI recognition sequence in pBD64 was eliminated by digestion with EcoRI followed by treatment with Klenow and deoxynucleotide triphosphates (Maniatis, T., et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold spring Harbor Laboratory, Cold Spring Harbor, N.Y.). Blunt end ligation and transformation yielded pB0154. The unique AvaI recognition sequence in pBO154 was eliminated in a similar manner to yield pBO171. pB0171 was digested with BamHI and PvuII and treated with Klenow and deoxynucleotide triphosphates to create blunt ends. The 6.4 kb fragment was purified, ligated and transformed into LE392 cells (Enquest, L.W., et al. (1977) J. Mol. Biol. 111, 97-120), to yield pB0172 which retains the unique BamHI site. To facilitate subcloning of subtilisin mutants, a unique and silent KpnI site starting at codon 166 was introduced into the subtilisin gene from pS4.5 (Wells, J.A., et al. (1983) Nucleic Acids Res., 11, 7911-7925) by site-directed mutagenesis. The KpnI+ plasmid was digested with EcoRI and treated with Klenow and deoxynucleotide triphosphates to create a blunt end. The Klenow was inactivated by heating for 20 min at 68°C, and the DNA was digested with BamHI. The 1.5 kb blunt EcoRI-BamHI fragment containing the entire subtilisin was ligated with the 5.8 kb NruI-BamHI from pB0172 to yield pBO180. The ligation of the blunt NruI end to the blunt EcoRI end recreated an EcoRI site. Proceeding clockwise around pB0180 from the EcoRI site at the 5' end of the subtilisin gene is the unique BamHI site at the 3' end of the subtilisin gene, the chloramphenicol and neomycin resistance genes and UB110 gram positive replication origin derived from pBD64, the ampicillin resistance gene and gram negative replication origin derived from pBR327.

B. Construction of Random Mutagenesis Library

The 1.5 kb EcoRI-BamHI fragment containing the B. amyloliquefaciens subtilisin gene (Wells et al., 1983) from pB0180 was cloned into M13mp11 to give M13mp11 SUBT essentially as previously described (Wells, J.A., et al. (1986) J. Biol. Chem., 261,6564-6570). Deoxyuridine containing template DNA was prepared according to Kunkel (Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA, 82 488-492). Uridine containing template DNA (Kunkel, 1985) was purified by CsCl density gradients (Maniatis, T. et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). A primer (AvaI$^-$) having the sequence

5' GAAAAAAGA$\overset{*}{C}$CCTAGCGTCGCTTA

ending at codon -11, was used to alter the unique AvaI recognition sequence within the subtilisin gene. (The asterisk denotes the mismatches from the wild-type sequence and underlined is the altered AvaI site.)

The 5' phosphorylated AvaI primer (~320 pmol) and ~40 pmol (~120μg) of uridine containing M13mp11 SUBT template in 1.88 ml of 53 mM NaCl, 7.4 mM MgCl2 and 7.4 mM Tris.HCl (pH 7.5) were annealed by heating to 90°C for 2 min. and cooling 15 min at 24°C (Fig. 31). Primer extension at 24°C was initiated by addition of 100μL containing 1 mM in all four deoxynucleotide triphosphates, and 20μl Klenow fragment (5 units/l). The extension reaction was stopped every 15 seconds over ten min by addition of 10μl 0.25 M EDTA (pH 8) to 50μl aliquots of the reaction mixture. Samples were pooled, phenol chlorophorm extracted and DNA was precipitated twice by addition of 2.5 vol 100% ethanol, and washed twice with 70% ethanol.

The pellet was dried, and redissolved in 0.4 ml 1 mM EDTA, 10 mM Tris (pH 8).

Misincorporation of α-thiodeoxynucleotides onto the 3' ends of the pool of randomly terminated template was carried out by incubating four 0.2 ml solutions each containing one-fourth of the randomly terminated template mixture (~20μg), 0.25 mM of a given α-thiodeoxynucleotide triphosphate, 100 units AMV polymerase, 50 mM KCL, 10 mM $MgCl_2$, 0.4 mM dithiothreitol, and 50 mM Tris (pH 8.3) (Champoux, J.J. (1984) Genetics, 2, 454-464). After incubation at 37 • C for 90 minutes, misincorporation reactions were sealed by incubation for five minutes at 37 • C with 50 mM all four deoxynucleotide triphosphates (pH 8), and 50 units AMV polymerase. Reactions were stopped by addition of 25 mM EDTA (final), and heated at 68 • C for ten min to inactivate AMV polymerase. After ethanol precipitation and resuspension, synthesis of closed circular heteroduplexes was carried out for two days at 14 • C under the same conditions used for the timed extension reactions above, except the reactions also contained 1000 units T4 DNA ligase, 0.5 mM ATP and 1 mM β-mercaptoethanol. Simultaneous restriction of each heteroduplex pool with KpnI, BamHI, and EcoRI confirmed that the extension reactions were nearly quantitative. Heteroduplex DNA in each reaction mixture was methylated by incubation with 80μM S-adenosylmethionine and 150 units dam methylase for 1 hour at 37 • C. Methylation reactions were stopped by heating at 68 • C for 15 min.

One-half of each of the four methylated heteroduplex reactions were transformed into 2.5 ml competent E. coli JM101 (Messing, J. (1979) Recombinant DNA Tech. Bull., 2, 43-48). The number of independent transformants from each of the four transformations ranged from 0.4-2.0 x 10^5. After growing out phage pools, RF DNA from each of the four transformations was isolated and purified by centrifugation through CsCl density gradients. Approximately 2μg of RF DNA from each of the four pools was digested with EcoRI, BamHI and AvaI. The 1.5 kb EcoRI-BamHI fragment (i.e., AvaI resistant) was purified on low gel temperature agarose and ligated into the 5.5 kb EcoRI-BamHI vector fragment of pB0180. The total number of independent transformants from each α-thiodeoxynucleotide misincorporation plasmid library ranged from 1.2-2.4 x 10^4. The pool of plasmids from each of the four transformations was grown out in 200 ml LB media containing 12.5μg/ml cmp and plasmid DNA was purified by centrifugation through CsCl density gradients.

C. Expression and Screening of Subtilisin Point Mutants

Plasmid DNA from each of the four misincorporation pools was transformed (Anagnostopoulos, C., et al. (1967), J. Bacteriol., 81, 741-746) into BG2036. For each transformation, 5μg of DNA produced approximately 2.5 x 10^5 independent BG2036 transformants, and liquid culture aliquots from the four libraries were stored in 10% glycerol at 70 • C. Thawed aliquots of frozen cultures were plated on LB/5μg/ml cmp/1.6% skim milk plates (Wells, J.A., et al. (1983) Nucleic Acids Res., 11, 7911-7925), and fresh colonies were arrayed onto 96-well microtiter plates containing 150 l per well LB media plus 12.5μg/ml cmp. After 1 h at room temperature, a replica was stamped (using a matched 96 prong stamp) onto a 132 mm BA 85 nitrocellulose filter (Schleicher and Scheull) which was layered on a 140 mm diameter LB/cmp/skim milk plate. Cells were grown about 16 h at 30 • C until halos of proteolysis were roughly 5-7 mm in diameter and filters were transferred directly to a freshly prepared agar plate at 37 • C containing only 1.6% skim milk and 50 mM sodium phosphate pH 11.5. Filters were incubated on plates for 3-6 h at 37 • C to produce halos of about 5 mm for wild-type subtilisin and were discarded. The plates were stained for 10 min at 24 • C with Coomassie blue solution (0.25% Coomassie blue (R-250) 25% ethanol) and destained with 25% ethanol, 10% acetic acid for 20 min. Zones of proteolysis appeared as blue halos on a white background on the underside of the plate and were compared to the original growth plate that was similarly stained and destained as a control. Clones were considered positive that produced proportionately larger zones of proteolysis on the high pH plates relative to the original growth plate. Negative clones gave smaller halos under alkaline conditions. Positive and negative clones were restreaked to colony purify and screened again in triplicate to confirm alkaline pH results.

D. Identification and Analysis of Mutant Subtilisins

Plasmid DNA from 5 ml overnight cultures of more alkaline active B.subtilis clones was prepared according to Birnboim and Doly (Birnboim, H.C., et al. (1979) Nucleic Acid Res. 7, 1513) except that incubation with 2 mg/ml lysozyme proceeded for 5 min at 37 • C to ensure cell lysis and an additional phenol/$CHCl_3$ extraction was employed to remove contaminants. The 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was ligated into M13mp11 and template DNA was prepared for DNA sequencing (Messing, J., et al. (1982) Gene, 19 269-276). Three DNA sequencing primers ending at codon 26, +95, and +155 were synthesized to match the subtilisin coding sequence. For preliminary sequence

identification a single track of DNA sequence, corresponding to the dNTPaS misincorporation library from which the mutant came, was applied over the entire mature protein coding sequence (i.e., a single dideoxyguanosine sequence track was applied to identify a mutant from the dGTPas library). A complete four track of DNA sequence was performed 200 bp over the site of mutagenesis to confirm and identify the mutant sequence (Sanger, F., et al., (1980) J. Mol. Biol., 143, 161-178). Confirmed positive and negative bacilli clones were cultured in LB media containing 12.5µg/mL cmp and purified from culture supernatants as previously described (Estell, D.A., et al. (1985) J. Biol. Chem., 260, 6518-6521). Enzymes were greater than 98% pure as analyzed by SDS-polyacrylamide gel electrophoresis (Laemmli, U.K. (1970), Nature, 227, 680-685), and protein concentrations were calculated from the absorbance at 280 nm,

$$\varepsilon_{280}^{0.1\%} = 1.17$$

(Maturbara, H., et al. (1965), J. Biol. Chem, 240, 1125-1130).

Enzyme activity was measured with 200µg/mL succinyl-L-AlaL-AlaL-ProL-Phep-nitroanilide (Sigma) in 0.1M Tris pH 8.6 or 0.1 M CAPS pH 10.8 at 25°C. Specific activity (µ moles product/min-mg) was calculated from the change in absorbance at 410 nm from production of p-nitroaniline with time per mg of enzyme (E410 = 8,480 M-lcm-l; Del Mar, E.G., et al. (1979), Anal. Biochem., 99, 316-320). Alkaline autolytic stability studies were performed on purified enzymes (200µg/mL) in 0.1 M potassium phosphate (pH 12.0) at 37°C. At various times aliquots were assayed for residual enzyme activity (Wells, J.A., et al. (1986) J. Biol. Chem., 261, 6564-6570).

E. Results

1. Optimization and analysis of mutagenesis frequency

A set of primer-template molecules that were randomly 3'-terminated over the subtilisin gene (Fig. 31) was produced by variable extension from a fixed 5'-primer (The primer mutated a unique AvaI site at codon 11 in the subtilisin gene). This was achieved by stopping polymerase reactions with EDTA after various times of extension. The extent and distribution of duplex formation over the 1 kb subtilisin gene fragment was assessed by multiple restriction digestion (not shown). For example, production of new HinfI fragments identified when polymerase extension had proceeded past Ile110, Leu233, and Asp259 in the subtilisin gene.

Misincorporation of each dNTPαs at randomly terminated 3' ends by AMV reverse transcriptase (Zakour, R.A., et al. (1982), Nature, 295, 708-710; Zakour, R.A., et al. (1984), Nucleic Acids Res., 12, 6615-6628) used conditions previously described (Champoux, J.J., (1984), Genetics, 2, 454-464). The efficiency of each misincorporation reaction was estimated to be greater than 80% by the addition of each dNTPαs to the AvaI restriction primer, and analysis by polyacrylamide gel electrophoresis. Misincorporations were sealed by polymerization with all four dNTP's and closed circular DNA was produced by reaction with DNA ligase.

Several manipulations were employed to maximize the yield of the mutant sequences in the heteroduplex. These included the use of a deoxyuridine containing template (Kunkel, T.A. (1985), Proc. Natl. Acad. Sci. USA, 82 488-492; Pukkila, P.J. et al. (1983), Genetics, 104, 571-582), in vitro methylation of the mutagenic strand (Kramer, W. et al. (1982) Nucleic Acids Res., 10 6475-6485), and the use of AvaI restriction-selection against the wild-type template strand which contained a unique AvaI site. The separate contribution of each of these enrichment procedures to the final mutagenesis frequency was not determined, except that prior to AvaI restriction-selection roughly one-third of the segregated clones in each of the four pools still retained a wild-type AvaI site within the subtilisin gene. After AvaI restriction-selection greater than 98% of the plasmids lacked the wild-type AvaI site.

The 1.5 kb EcoRI-BamHI subtilisin gene fragment that was resistant to AvaI restriction digestion, from each of the four CsCl purified M13 RF pools was isolated on low melting agarose. The fragment was ligated in situ from the agarose with a similarly cut E. coli-B. subtilis shuttle vector, pB0180, and transformed directly into E coli LE392. Such direct ligation and transformation of DNA isolated from agarose avoided loses and allowed large numbers of recombinants to be obtained (>100,000 per µg equivalent of input M13 pool).

The frequency of mutagenesis for each of the four dNTPαs misincorporation reactions was estimated from the frequency that unique restriction sites were eliminated (Table XX). The unique restriction sites

chosen for this analysis, ClaI, PvuII, and KpnI, were distributed over the subtilisin gene starting at codons 35, 104, and 166, respectively. As a control, the mutagenesis frequency was determined at the PstI site located in the β lactamase gene which was outside the window of mutagenesis. Because the absolute mutagenesis frequency was close to the percentage of undigested plasmid DNA, two rounds of restriction-selection were necessary to reduce the background of surviving uncut wild-type plasmid DNA below the mutant plasmid (Table XX). The background of surviving plasmid from wild-type DNA probably represents the sum total of spontaneous mutations, uncut wild-type plasmid, plus the efficiency with which linear DNA can transform E. coli. Subtracting the frequency for unmutagenized DNA (background) from the frequency for mutant DNA, and normalizing for the window of mutagenesis sampled by a given restriction analysis (4-6 bp) provides an estimate of the mutagenesis efficiency over the entire coding sequence (~1000 bp).

TABLE XX

α-thiol dNTP misincorporated[b]	Restriction Site Selection	% resistant clones[c] 1st round	2nd round	Total	% resistant clones over Background[d]	% mutants per 1000bp[e]
None	PstI	0.32	0.7	0.002	0	-
G	PstI	0.33	1.0	0.003	0.001	0.2
T	PstI	0.32	<0.5	<0.002	0	0
C	PstI	0.43	3.0	0.013	0.011	3
None	ClaI	0.28	5	0.014	0	-
G	ClaI	2.26	85	1.92	1.91	380
T	ClaI	0.48	31	0.15	0.14	35
C	ClaI	0.55	15	0.08	0.066	17
None	PvuII	0.08	29	0.023	0	-
G	PvuII	0.41	90	0.37	0.35	88
T	PvuII	0.10	67	0.067	0.044	9
C	PvuII	0.76	53	0.40	0.38	95
None	KpnI	0.41	3	0.012	0	-
G	KpnI	0.98	35	0.34	0.33	83
T	KpnI	0.36	15	0.054	0.042	8
C	KpnI	1.47	26	0.38	0.37	93

(a) Mutagenesis frequency is estimated from the frequency for obtaining mutations that alter unique restriction sites within the mutagenized subtilisin gene (i.e., ClaI, PvuII, or KpnI) compared to mutation frequencies of the PstI site, that is outside the window of mutagenesis.

(b) Plasmid DNA was from wild-type (none) or mutagenized by dNTPαs misincorporation as described.

(c) Percentage of resistant clones was calculated from the fraction of clones obtained after three fold or greater over-digestion of the plasmid with the indicated restriction enzyme compared to a

non-digested control. Restriction-resistant plasmid DNA from the first round was subjected to a second round of restriction-selection. The total represents the product of the fractions of resistant clones obtained from both rounds of selection and gives percentage of restriction-site mutant clones in the original starting pool. Frequencies were derived from counting at least 20 colonies and usually greater than 100.

(d) Percent resistant clones was calculated by subtracting the percentage of restriction-resistant clones obtained for wild-type DNA (i.e., none) from that obtained for mutant DNA.

(e) This extrapolates from the frequency of mutation over each restriction site to the entire subtilisin gene (~1 kb). This has been normalized to the number of possible bases (4-6 bp) within each restriction site that can be mutagenized by a given misincorporation event.

From this analysis, the average percentage of subtilisin genes containing mutations that result from dGTPαs, dCTPαs, or dTTPαs misincorporation was estimated to be 90, 70, and 20 percent, respectively. These high mutagenesis frequencies were generally quite variable depending upon the dNTPαs and misincorporation efficiencies at this site. Misincorporation efficiency has been reported to be both dependent on the kind of mismatch, and the context of primer (Champoux, J.J., (1984); Skinner, J.A., et al. (1986) Nucleic Acids Res., 14, 6945-6964). Biased misincorporation efficiency of dGTPαs and dCTPαs over dTTPαs has been previously observed (Shortle, D., et al. (1985), Genetics, 110, 539-555). Unlike the dGTPαs, dCTPαs, and dTTPαs libraries the efficiency of mutagenesis for the dATPαs misincorporation library could not be accurately assessed because 90% of the restriction-resistant plasmids analyzed simply lacked the subtilisin gene insert. This problem probably arose from self-ligation of the vector when the dATPαs mutagenized subtilisin gene was subcloned from M13 into pB0180. Correcting for the vector background, we estimate the mutagenesis frequency around 20 percent in the dATPαs misincorporation library. In a separate experiment (not shown), the mutagenesis efficiencies for dGTPαs and dTTPαs misincorporation were estimated to be around 50 and 30 percent, respectively, based on the frequency of reversion of an inactivating mutation at codon 169.

The location and identity of each mutation was determined by a single track of DNA sequencing corresponding to the misincorporated αthiodeoxynucleotide over the entire gene followed by a complete four track of DNA sequencing focused over the site of mutation. Of 14 mutants identified, the distribution was similar to that reported by Shortle and Lin (1985) except we did not observe nucleotide insertion or deletion mutations. The proportion of AG mutations was highest in the G misincorporation library, and some unexpected point mutations appeared in the dTTPαs and dCTPαs libraries.

2. Screening and Identification of Alkaline Stability Mutants of Subtilisin

It is possible to screen colonies producing subtilisin by halos of casein digestion (Wells, J.A. et al. (1983) Nucleic Acids Res., 11, 7911-7925). However, two problems were posed by screening colonies under high alkaline conditions (>pH 11). First, B. subtilis will not grow at high pH, and we have been unable to transform an alkylophilic strain of bacillus. This problem was overcome by adopting a replica plating strategy in which colonies were grown on filters at neutral pH to produce subtilisin and filters subsequently transferred to casein plates at pH 11.5 to assay subtilisin activity. However, at pH 11.5 the casein micells no longer formed a turbid background and thus prevented a clear observation of proteolysis halos. The problem was overcome by briefly staining the plate with Coomassie blue to amplify proteolysis zones and acidifying the plates to develop casein micell turbidity. By comparison of the halo size produced on the reference growth plate (pH 7) to the high pH plate (pH 11.5), it was possible to identify mutant subtilisins that had increased (positives) or decreased (negatives) stability under alkaline conditions.

Roughly 1000 colonies were screened from each of the four misincorporation libraries. The percentage of colonies showing a differential loss of activity at pH 11.5 versus pH 7 represented 1.4, 1.8, 1.4, and 0.6% of the total colonies screened from the thiol dGTPαs, dATPαs, dTTPαs, and dCTPαs libraries, respectively. Several of these negative clones were sequenced and all were found to contain a single base change as expected from the misincorporation library from which they came. Negative mutants included A36, E170 and V50. Two positive mutants were identified as V107 and R213. The ratio of negatives to positives was roughly 50:1.

3. Stability and Activity of Subtilisin Mutants at Alkaline pH

Subtilisin mutants were purified and their autolytic stabilities were measured by the time course of inactivation at pH 12.0 (Figs. 32 and 33). Positive mutants identified from the screen (i.e., V107 and R213) were more resistant to alkaline induced autolytic inactivation compared to wild-type; negative mutants (i.e., E170 and V50) were less resistant. We had advantageously produced another mutant at position 50 (F50) by site-directed mutagenesis. This mutant was more stable than wild-type enzyme to alkaline autolytic inactivation (Fig. 33) At the termination of the autolysis study, SDS-PAGE analysis confirmed that each subtilisin variant had autolyzed to an extent consistent with the remaining enzyme activity.

The stabilizing effects of V107, R213, and F50 are cumulative. See Table XXI. The double mutant, V107/R213 (made by subcloning the 920 bp EcoRI-KpnI fragment of pB0180V107 into the 6.6 kb EcoRI-KpnI fragment of pB0180R213), is more stable than either single mutant. The triple mutant, F50/V107/R213 (made by subcloning the 735 bp EcoRI-PvuII fragment of pF50 (Example 2) into the 6.8 kb EcoRI-PvuII fragment of pB0180/V107, is more stable than the double mutant V107/R213 or F50. The inactivation curves show a biphasic character that becomes more pronounced the more stable the mutant analyzed. This may result from some destablizing chemical modification(s) (eg., deamidation) during the autolysis study and/or reduced stabilization caused by complete digestion of larger autolysis peptides. These alkaline autolysis studies have been repeated on separately purified enzyme batches with essentially the same results. Rates of autolysis should depend both on the conformational stability as well as the specific activity of the subtilisin variant (Wells, J.A., et al. (1986), J. Biol. Chem., 261, 6564-6570). It was therefore possible that the decreases in autolytic inactivation rates may result from decreases in specific activity of the more stable mutant under alkaline conditions. In general the opposite appears to be the case. The more stable mutants, if anything, have a relatively higher specific activity than wild-type under alkaline conditions and the less stable mutants have a relatively lower specific activity. These subtle effects on specific activity for V107/R213 and F50/V107/R213 are cumulative at both pH 8.6 and 10.8. The changes in specific activity may reflect slight differences in substrate specificity, however, it is noteworthy that only positions 170 and 107 are within 6A of a bound model substrate (Robertus, J.D., et al. (1972), Biochemistry 11, 2438-2449).

TABLE XXI

Relationship between relative specific acitivity at pH 8.6 or 10.8 and alkaline autolytic stability			
Enzyme	Relative specific activity		Alkaline autolysis half-time (min)b
	pH 8.6	pH 10.8	
Wild-type	100±1	100±3	86
Q170	46±1	28±2	13
V107	126±3	99±5	102
R213	97±1	102±1	115
V107/R213	116±2	106±3	130
V50	66±4	61±1	58
F50	123±3	157±7	131
F50/V107/R213	126±2	152±3	168

[a] Relative specific activity was the average from triplicate activity determinations divided by the wild-type value at the same pH. The average specific activity of wild-type enzyme at pH 8.6 and 10.8 was 70μmoles/min-mg and 37μmoles/min-mg, respectively.
[b] Time to reach 50% activity was taken from Figs. 32 and 33.

F. Random Cassette Mutagenesis of Residues 197 through 228

Plasmid pΔ222 (Wells, et al. (1985) Gene 34, 315-323) was digested with PstI and BamHI and the 0.4 kb PstI/BamHI fragment (fragment 1, see Fig. 34) purified from a polyacrylamide gel by electroelution.

The 1.5 kb EcoRI/BamHI fragment from pS4.5 was cloned into M13mp9. Site directed mutagenesis was performed to create the A197 mutant and simultaneously insert a silent SstI site over codons 195-196. The mutant EcoRI/BamHI fragment was cloned back into pBS42. The pA197 plasmid was digested with BamHI and SstI and the 5.3 kb BamHI/SstI fragment (fragment 2) was purified from low melting agarose.

Complimentary oligonucleotides were synthesized to span the region from SstI (codons 195-196) to PstI (codons 228-230). These oligodeoxynucleotides were designed to (1) restore codon 197 to the wild type, (2) re-create a silent KpnI site present in pΔ222 at codons 219-220, (3) create a silent SmaI site over codons 210-211, and (4) eliminate the PstI site over codons 228-230 (see Fig. 35). Oligodeoxynucleotides were synthesized with 2% contaminating nucleotides at each cycle of synthesis, e.g., dATP reagent was spiked with 2% dCTP, 2% dGTP, and 2% dTTP. For 97-mers, this 2% poisoning should give the following percentages of non-mutant, single mutants and double or higher mutants per strand with two or more misincorporations per complimentary strand: 14% non-mutant, 28% single mutant, and 57% with ≥2 mutations, according to the general formula

$$f = \frac{\mu^n}{n!} e^{-\mu} .$$

where μ is the average number of mutations and n is a number class of mutations and f is the fraction of the total having that number of mutations. Complimentary oligodeoxynucleotide pools were phosphorylated and annealed (fragment 3) and then ligated at 2-fold molar excess over fragments 1 and 2 in a three-way ligation.

E. coli MM294 was transformed with the ligation reaction, the transformation pool-grown up over night and the pooled plasmid DNA was isolated. This pool represented 3.4×10^4 independent transformants. This plasmid pool was digested with PstI and then used to retransform E. coli. A second plasmid pool was prepared and used to transform B. subtilis (BG2036). Approximately 40% of the BG2036 transformants actively expressed subtilisin as judged by halo-clearing on casein plates. Several of the non-expressing transformants were sequenced and found to have insertions or deletions in the synthetic cassettes. Expressing BG2036 mutants were arrayed in microtiter dishes with 150μl of LB/12.5μg/mL chloramphenicol (cmp) per well, incubated at 37 • C for 3-4 hours and then stamped in duplicate onto nitrocellulose filters laid on LB 1.5% skim milk/5μg/mL cmp plates and incubated overnight at 33 • C (until halos were approximately 4-8 mm in diameter). Filters were then lifted to stacks of filter paper saturated with 1 x Tide commercial grade detergent, 50 mM Na_2CO_3, pH 11.5 and incubated at 65 • C for 90 min. Overnight growth plates were Commassie stained and destained to establish basal levels of expression. After this treatment, filters were returned to pH7/skim milk/20μg/mL tetracycline plates and incubated at 37 • C for 4 hours to overnight.

Mutants identified by the high pH stability screen to be more alkaline stable were purified and analyzed for autolytic stability at high pH or high temperature. The double mutant C204/R213 was more stable than wild type at either high pH or high temperature (Table XXII).

This mutant was dissected into single mutant parents (C204 and R213) by cutting at the unique SmaI restriction site (Fig. 35) and either ligating wild type sequence 3' to the SmaI site to create the single C204 mutant or ligating wild type sequence 5' to the SmaI site to create the single R213 mutant. Of the two single parents, C204 was nearly as alkaline stable as the parent double mutant (C04/R213) and slightly more thermally stable. See Table XXII. The R213 mutant was only slightly more stable than wild type under both conditions (not shown).

Another mutant identified from the screen of the 197 to 228 random cassette mutagenesis was R204. This mutant was more stable than wild type at both high pH and high temperature but less stable than C204.

TABLE XXII

Stability of subtilisin variants

Purified enzymes (200μg/mL) were incubated in 0.1M phosphate, pH 12 at 30°C for alkaline autolysis, or in 2mM $CaCl_2$, 50mM MOPS, pH 7.0 at 62°C for thermal autolysis. At various times samples were assayed for residual enzyme activity. Inactivations were roughly pseudo-first order, and t 1/2 gives the time it took to reach 50% of the starting activity in two separate experiments.

	t 1/2 (alkaline autolysis)		t 1/2 (thermal autolysis)	
Subtilisin variant	Exp. #1	Exp. #2	Exp. #1	Exp. #2
wild type	30	25	20	23
F50/V107/R213	49	41	18	23
R204	35	32	24	27
C204	43	46	38	40
C204/R213	50	52	32	36
L204/R213	32	30	20	21

G. Random Mutagenesis at Codon 204

Based on the above results, codon 204 was targeted for random mutagenesis. Mutagenic DNA cassettes (for codon at 204) all contained a fixed R213 mutation which was found to slightly augment the stability of the C204 mutant.

Plasmid DNA encoding the subtilisin mutant C204/R213 was digested with SstI and EcoRI and a 1.0 kb EcoRI/SstI fragment was isolated by electro-elution from polyacrylamide gel (fragment 1, see Fig. 35).

C204/R213 was also digested with SmaI and EcoRI and the large 4.7 kb fragment, including vector sequences and the 3' portion of coding region, was isolated from low melting agarose (fragment 2, see Fig. 36).

Fragments 1 and 2 were combined in four separate three-way ligations with heterophosphorylated fragments 3 (see Figs. 36 and 37). This heterophosphorylation of synthetic duplexes should preferentially drive the phosphorylated strand into the plasmid ligation product. Four plasmid pools, corresponding to the four ligations, were restricted with SmaI in order to linearize any single cut C204/R213 present from fragment 2 isolation, thus reducing the background of C204/R213. E. coli was then re-transformed with

SmaI-restricted plasmid pools to yield a second set of plasmid pools which are essentially free of C204/R213 and any non-segregated heterduplex material.

These second enriched plasmid pools were then used to transform B. subtilis (BG2036) and the resulting four mutant pools were screened for clones expressing subtilisin resistant to high pH/temperature inactivation. Mutants found positive by such a screen were further characterized and identified by sequencing.

The mutant L204/R213 was found to be slightly more stable than the wild type subtilisin. See Table XXII.

Having described the preferred embodiments of the present invention, it will appear to those ordinarily skilled in the art that various modifications may be made to the disclosed embodiments, and that such modifications are intended to be within the scope of the present invention.

Claims

1. A subtilisin mutant derived by the substitution of at least one amino acid residue of a precursor subtilisin with a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterised by the substitution at one or more of Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of Bacillus amyloliquefaciens subtilisin and equivalent amino acid residues in other precursor subtilisins.

2. A subtilisin mutant having an amino acid sequence derived from the amino acid sequence of a precursor subtilisin by the substitution of more than one amino acid residue of said amino acid sequence of said precursor subtilisin by a different amino acid, so that the subtilisin mutant has at least one property which is different from the same property of the precursor subtilisin, characterized by substitutions at more than one of Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214, and Gly215 of Bacillus amyloliquefaciens subtilisin and equivalent amino acid residues in other precursor subtilisins, with the proviso that when substitution is made at any residue in the group Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156 Gly166, Gly169, Phe189, Tyr217 and Met222 a substitution is also made at at least one specified position not of that group.

3. The mutant of claim 2 wherein said combinations are selected from Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 and Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.

4. A subtilisin mutant derived by the deletion of one or more amino acid residues in a precursor subtilisin equivalent to 161-164 in B. amyloliquefaciens subtilisin, said deletion being made alone or in combination with substitutions in the amino acid sequence of the precursor subtilisin, and producing at least one property which is different from the same property of the precursor subtilisin.

5. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Leu + 126 of B. amyloliquefaciens subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.

6. A subtilisin mutant having altered substrate specificity when compared to a precursor subtilisin, the mutant being derived by the substitution of a different amino acid at the residue equivalent to Asp + 99 in B. amyloliquefaciens subtilisin, alone or in combination with other substitutions or deletions in the amino acid sequence of the precursor subtilisin.

7. A DNA sequence encoding the mutant of any one of the preceding claims.

8. An expression vector containing the mutant DNA sequence of claim 7.

9. A host cell transformed with the expression vector or claim 8.

Patentansprüche

1. Subtilisinmutante, die durch Substitution zumindest eines Aminosäurerests eines Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft aufweist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet, gekennzeichnet durch die Substitution an einem oder mehreren von Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von Bacillus amyloliquefaciens-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen.

2. Subtilisinmutante mit einer Aminosäuresequenz, die aus der Aminosäuresequenz eines Vorläufer-Subtilisins durch Substitution mehr als eines Aminosäurerests der Aminosäuresequenz des Vorläufer-Subtilisins durch eine davon verschiedene Aminosäure hergeleitet ist, sodaß die Subtilisinmutante zumindest eine Eigenschaft auWeist, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet, gekennzeichnet durch Substitutionen an mehr als einem von Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 und Gly215 von Bacillus amyloliquefaciens-Subtilisin und äquivalenten Aminosäureresten in anderen Vorläufer-Subtilisinen, mit der Maßgabe, daß bei einer Substitution an irgendeinem Rest in der Gruppe Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 und Met222 eine Substitution auch an zumindest einer bestimmten Position durchgeführt wird, die nicht dieser Gruppe angehört.

3. Mutante nach Anspruch 2, worin die Kombinationen aus Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Tyr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 und Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217 ausgewählt sind.

4. Subtilisinmutante, die durch Löschung eines oder mehrerer Aminosäurereste in einem Vorläufer-Subtilisin, das 161-164 in B. amyloliquefaciens-Subtilisin äquivalent ist, hergeleitet ist, wobei die Löschung entweder alleine oder in Kombination mit Substitutionen in der Aminosäuresequenz des Vorläufer-Subtilisins erfolgt, und zumindest eine Eigenschaft ergibt, die sich von der gleichen Eigenschaft des Vorläufer-Subtilisins unterscheidet.

5. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufersubtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Leu + 126 von B. amyloliquefaciens-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.

6. Subtilisinmutante mit geänderter Substratspezifität im Vergleich zu einem Vorläufersubtilisin, wobei die Mutante durch Substitution einer unterschiedlichen Aminosäure am Rest, der Asp +99 im B. amyloliquefaciens-Subtilisin äquivalent ist, alleine oder in Kombination mit anderen Substitutionen oder Löschungen in der Aminosäuresequenz des Vorläufer-Subtilisins hergeleitet ist.

7. DNA-Sequenz, die für die Mutante nach einem der vorhergehenden Ansprüche kodiert.

8. Expressionsvektor, der die Mutanten-DNA-Sequenz von Anspruch 7 enthält.

9. Wirtszelle, die mit dem Expressionsvektor von Anspruch 8 transformiert ist.

Revendications

1. Mutant de subtilisine dérivé par la substitution d'au moins un résidu d'acide aminé d'une subtilisine précurseur et par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par la substitution à un ou plusieurs de Tyr21, Thr22, Ser24, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Ile107, Gly110, Met124, Lys170, Tyr171, Pro172, Asp197, Met199, Ser204, Lys213, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilise de Bacillus amyloliquefaciens et les résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs.

2. Mutant de subtilisine ayant une séquence d'acides aminés dérivée de la séquence d'acides aminés d'une subtilisine précurseur par la substitution de plus d'un résidu d'acide aminé de ladite séquence d'acides aminés de ladite subtilisine précurseur par un acide aminé différent de manière que le mutant de subtilisine ait au moins une propriété qui est différente de la même propriété de la subtilisine précurseur, caractérisé par des substitutions à plus d'un de Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Ala48, Ser49, Met50, Ser87, Lys94, Val95, Tyr104, Ile107, Gly110, Met124, Ala152, Asn155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu135, Gly97, Ser101, Gly102, Glu103, Gly127, Gly128, Pro129, Tyr214 et Gly215 de la subtilisine de Bacillus amyloliquefaciens et des résidus d'acides aminés équivalents dans d'autres subtilisines précurseurs, à condition que quand la substitution est effectuée à tout résidu dans le groupe formé de Asp32, Ser33, Tyr104, Ala152, Asn155, Glu156, Gly166, Gly169, Phe189, Tyr217 et Met222, une substitution soit également effectuée en au moins une position spécifiée ne faisant pas partie de ce groupe.

3. Mutant de la revendication 2 où lesdites associations sont choisies parmi Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 et Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.

4. Mutant de subtilisine dérivé par la délétion d'un ou plusieurs résidus d'acides aminés dans une subtilisine précurseur équivalente à 161-164 dans la subtilisine de B. amyloliquefaciens, ladite délétion étant effectuée seule ou en association avec des substitutions dans la séquence d'acides aminés de la subtilisine précurseur et la production d'au moins une propriété qui est différente de la même propriété de la subtilisine précurseur.

5. Mutant de subtilisine ayant une spécificité modifiée du substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Leu + 126 de la subtilisine de B. amyloliquefaciens, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.

6. Mutant de subtilisine ayant une spécificité modifiée de substrat en comparaison avec une subtilisine précurseur, le mutant étant dérivé par la substitution d'un acide aminé différent au résidu équivalent à Asp + 99 dans la substilisine de B. amyloliquefaciens, seule ou en association avec d'autres substitutions ou délétions dans la séquence d'acides aminés de la subtilisine précurseur.

7. Séquence d'ADN codant le mutant selon l'une quelconque des revendications précédentes.

8. Vecteur d'expression contenant la séquence d'ADN du mutant de la revendication 7.

9. Cellule hôte transformée par le vecteur d'expression de la revendication.8 .

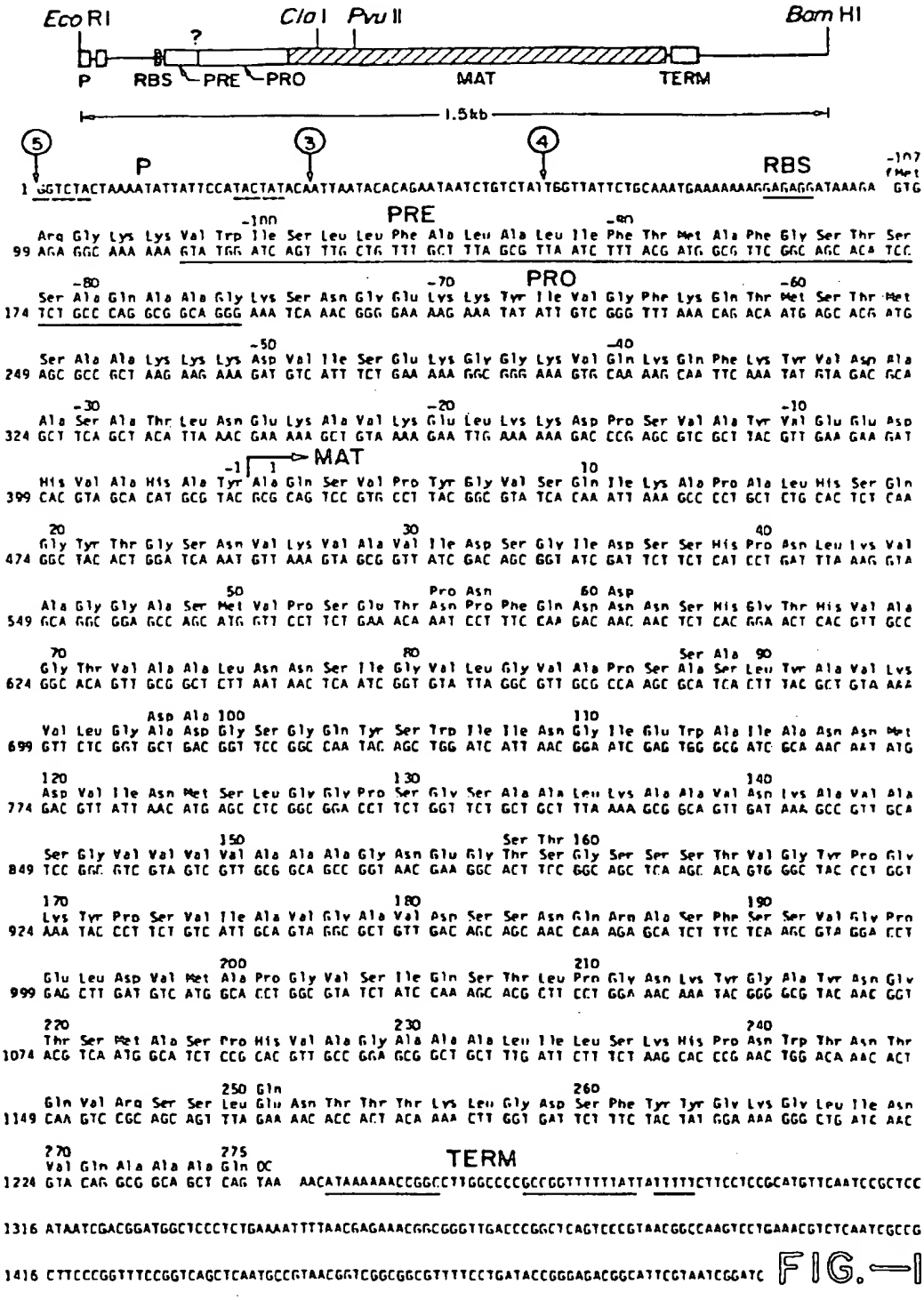

EcoRI
ClaI
PvuII
BamHI
?
P
RBS
PRE
PRO
MAT
TERM
1.5kb
5
3
4
P
RBS
-107
fMet
PRE
PRO
MAT
TERM

FIG.—1

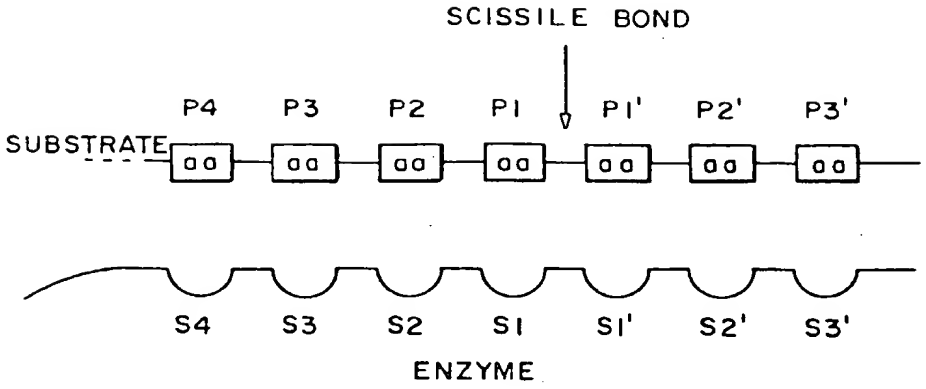

SCISSILE BOND
P4
P3
P2
P1
P1'
P2'
P3'
SUBSTRATE
S4
S3
S2
S1
S1'
S2'
S3'
ENZYME

FIG.—2

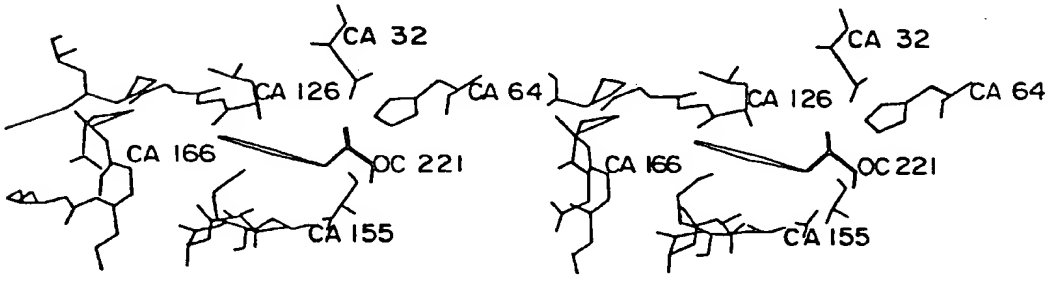

CA 32
CA 126
CA 64
CA 166
OC 221
CA 155
CA 32
CA 126
CA 64
CA 166
OC 221
CA 155

FIG.—3

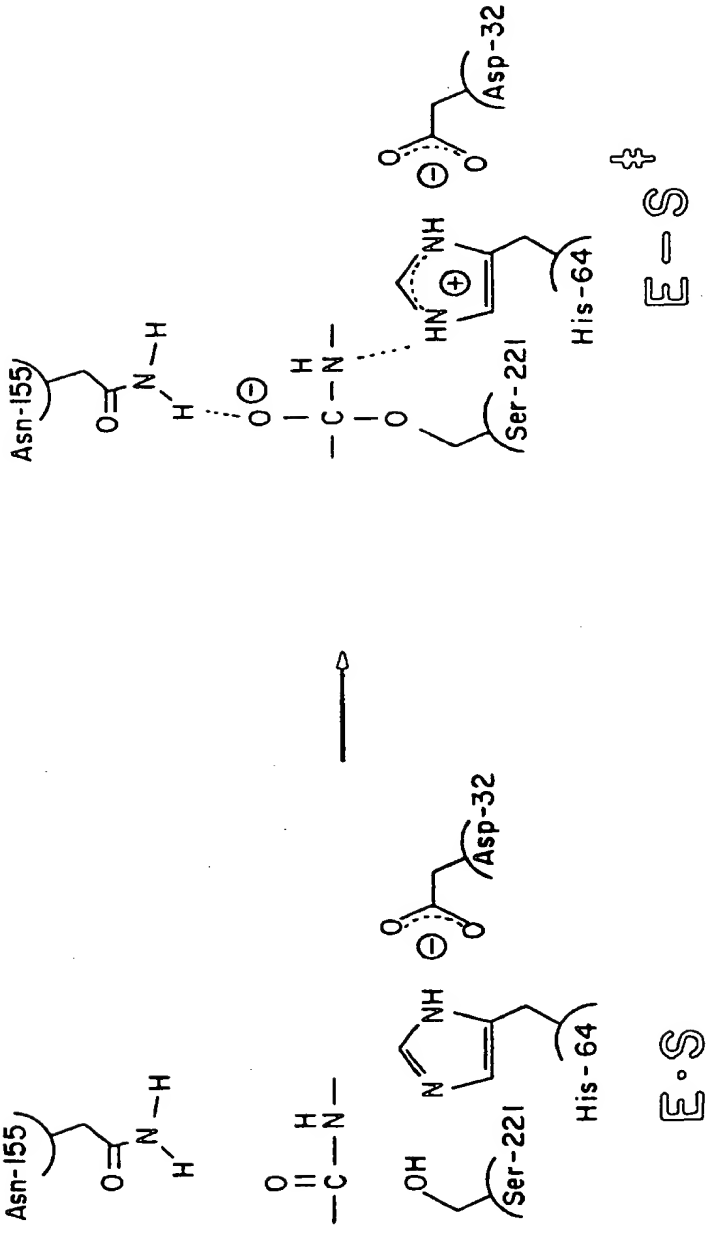


FIG.-4

Homology of Bacillus proteases

1.Bacillus amyloliquifaciens
2.Bacillus subtilis var.I168
3.Bacillus licheniformis (carlsbergensis)

1									10									20
A	Q	S	V	P	Y	G	V	S	Q	I	K	A	P	A	L	H	S	Q	G
A	Q	S	V	P	Y	G	I	S	Q	I	K	A	P	A	L	H	S	Q	G
A	Q	T	V	P	Y	G	I	P	L	I	K	A	D	K	V	Q	A	Q	G

21									30									40
Y	T	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P
Y	T	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P
F	K	G	A	N	V	K	V	A	V	L	D	T	G	I	Q	A	S	H	P

41									50									60
D	L	K	V	A	G	G	A	S	H	V	P	S	E	T	N	P	F	Q	D
D	L	N	V	R	G	G	A	S	F	V	P	S	E	T	N	P	Y	Q	D
D	L	N	V	V	G	G	A	S	F	V	A	G	E	A	Y	N	T	o	D

61									70									80
N	N	S	H	G	T	H	V	A	G	T	V	A	A	L	N	N	S	I	G
G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G
G	N	G	H	G	T	H	V	A	G	T	V	A	A	L	D	N	T	T	G

81									90									100
V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	G	A	D	G
V	L	G	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
V	L	G	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S	G

101									110									120
S	G	Q	Y	S	U	I	I	N	G	I	E	U	A	I	A	N	N	M	D
S	G	Q	Y	S	U	I	I	N	G	I	E	U	A	I	S	N	N	M	D
S	G	S	Y	S	G	I	V	S	G	I	E	U	A	T	T	N	G	M	D

FIG.—5A—1

121									130									140
V	I	N	H	S	L	G	G	P	S	G	S	A	A	L	K	A	A	V	D
V	I	N	H	S	L	G	G	P	T	G	S	T	A	L	K	T	V	V	D
V	I	N	H	S	L	G	G	A	S	G	S	T	A	H	K	Q	A	V	D

141									150									160
K	A	V	A	S	G	V	V	V	V	A	A	A	G	N	E	G	T	S	G
K	A	V	S	S	G	I	V	V	A	A	A	A	G	N	E	G	S	S	G
N	A	Y	A	R	G	V	V	V	V	A	A	A	G	N	S	G	N	S	G

161									170									180
S	S	S	T	V	G	Y	P	G	K	Y	P	S	V	I	A	V	G	A	V
S	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V
S	T	N	T	I	G	Y	P	A	K	Y	D	S	V	I	A	V	G	A	V

181									190									200
D	S	S	N	Q	R	A	S	F	S	S	V	G	P	E	L	D	V	H	A
N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L	D	V	H	A
D	S	N	S	N	R	A	S	F	S	S	V	G	A	E	L	E	V	H	A

201									210									220
P	G	V	S	I	Q	S	T	L	P	G	N	K	Y	G	A	Y	N	G	T
P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T
P	G	A	G	V	Y	S	T	Y	P	T	N	T	Y	A	T	L	N	G	T

221									230									240
S	H	A	S	P	H	V	A	G	A	A	A	L	I	L	S	K	H	P	N
S	H	A	T	P	H	V	A	G	A	A	A	L	I	L	S	K	H	P	T
S	H	A	S	P	H	V	A	G	A	A	A	L	I	L	S	K	H	P	N

241									250									260
U	T	N	T	Q	V	R	S	S	L	E	N	T	T	T	K	L	G	D	S
U	T	N	A	Q	V	R	D	R	L	E	S	T	A	T	Y	L	G	N	S
L	S	A	S	Q	V	R	N	R	L	S	S	T	A	T	Y	L	G	S	S

261									270					
F	Y	Y	G	K	G	L	I	N	V	Q	A	A	A	Q
F	Y	Y	G	K	G	L	I	N	V	Q	A	A	A	Q
F	Y	Y	G	K	G	L	I	N	V	E	A	A	A	Q

FIG.— 5A—2

ALIGNMENT OF B.AMYLOLIQUIFACIENS SUBTILISIN AND THERMITASE
1.B.amyloliquifaciens subtilisin
2.thermitase

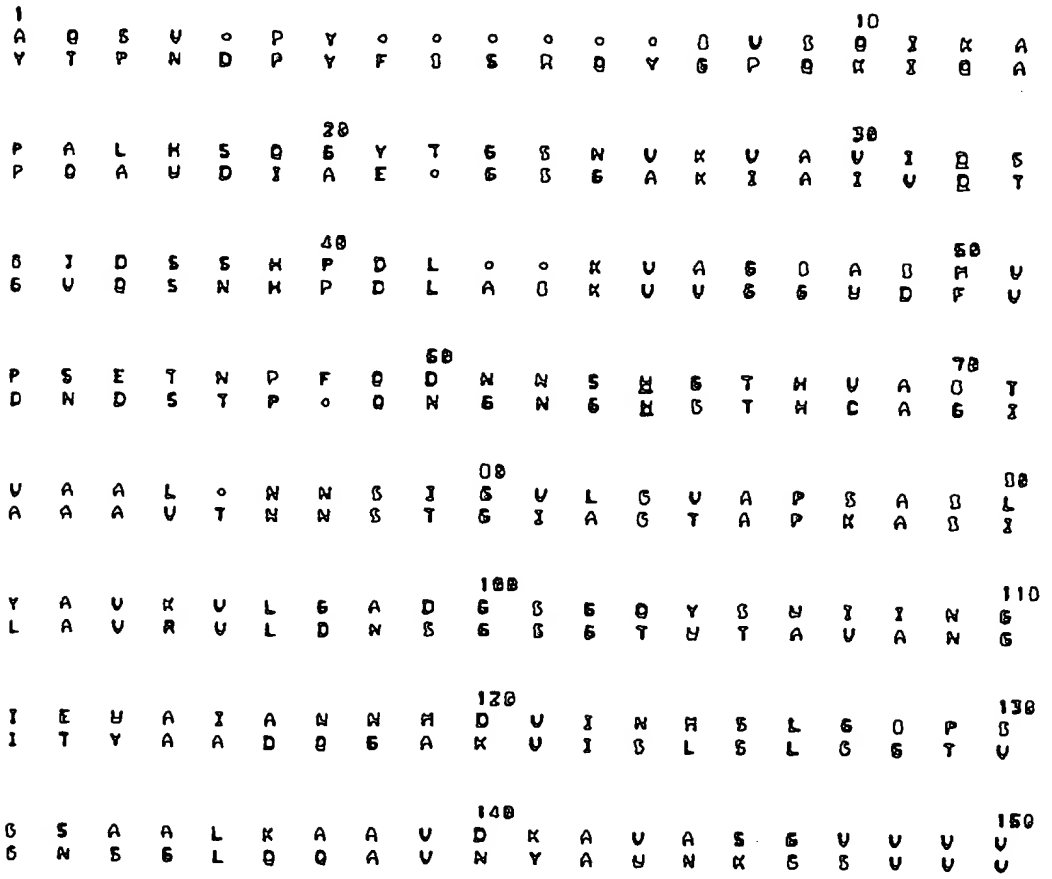


FIG.—5B—1

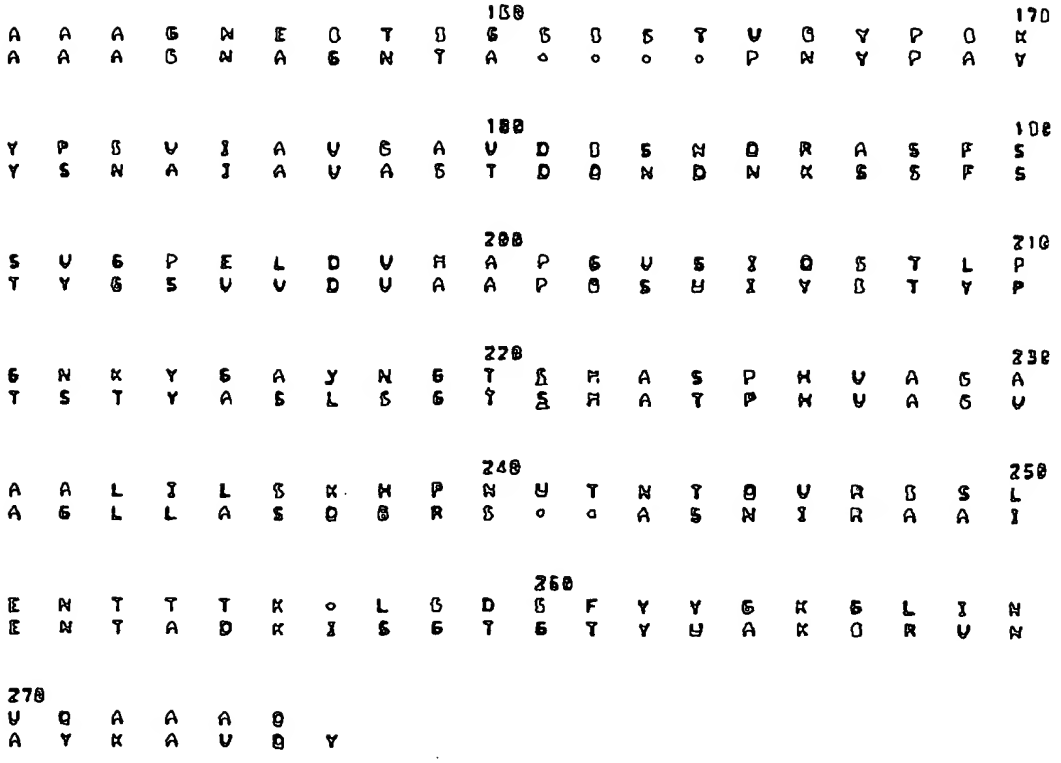


FIG.—5B-2

TOTALLY CONSERVED RESIDUES IN SUBTILISINS

```
1                                  10                                  20
.  .  .  .  P  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

21                                 30                                  40
.  .  G  .  .  .  .  .  .  .  .  .  D  .  G  .  .  .  .  H  .  .  .  .

41                                 50                                  60
.  .  .  .  .  G  .  .  .  .  V  .  .  .  .  .  .  .  .  .  .  .  .

61                                 70                                  80
.  .  .  H  G  T  H  .  .  G  .  .  .  .  .  .  .  .  .  .  .  .  .

81                                 90                                  100
.  .  G  .  .  .  .  .  .  .  .  .  .  .  .  .  V  L  .  .  .  G  .

101                                110                                 120
S  G  .  .  .  .  .  .  .  .  G  .  .  .  .  .  .  .  .  .  .  .  .

121                                130                                 140
.  .  .  .  .  L  G  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

141                                150                                 160
.  .  .  .  .  G  .  .  .  .  .  .  .  .  G  N  .  .  .  .  .  .  .

161                                170                                 180
.  .  .  .  .  .  Y  P  .  .  .  .  .  .  .  .  V  .  .  .  .  .  .

181                                190                                 200
.  .  .  .  .  .  .  S  F  S  .  .  .  .  .  .  .  .  .  .  .  .  .

201                                210                                 220
P  G  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  G  T  .

221                                230                                 240
S  M  A  .  P  H  V  A  G  .  .  .  .  .  .  .  .  .  .  .  .  .  .

241                                250                                 260
.  .  .  .  .  .  R  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

261                                270
.  .  .  .  .  .  .  .  N  .  .  .  .  .  .
```

FIG.—5C

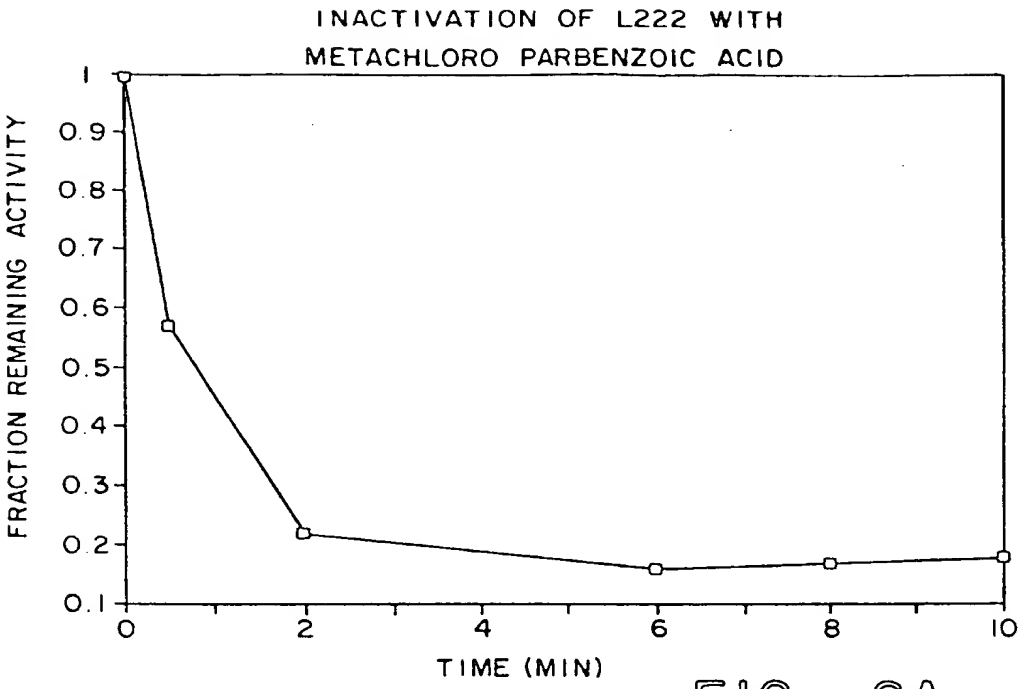

INACTIVATION OF L222 WITH
METACHLORO PARBENZOIC ACID
FRACTION REMAINING ACTIVITY
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
4
6
8
10
TIME (MIN)

FIG.—6A

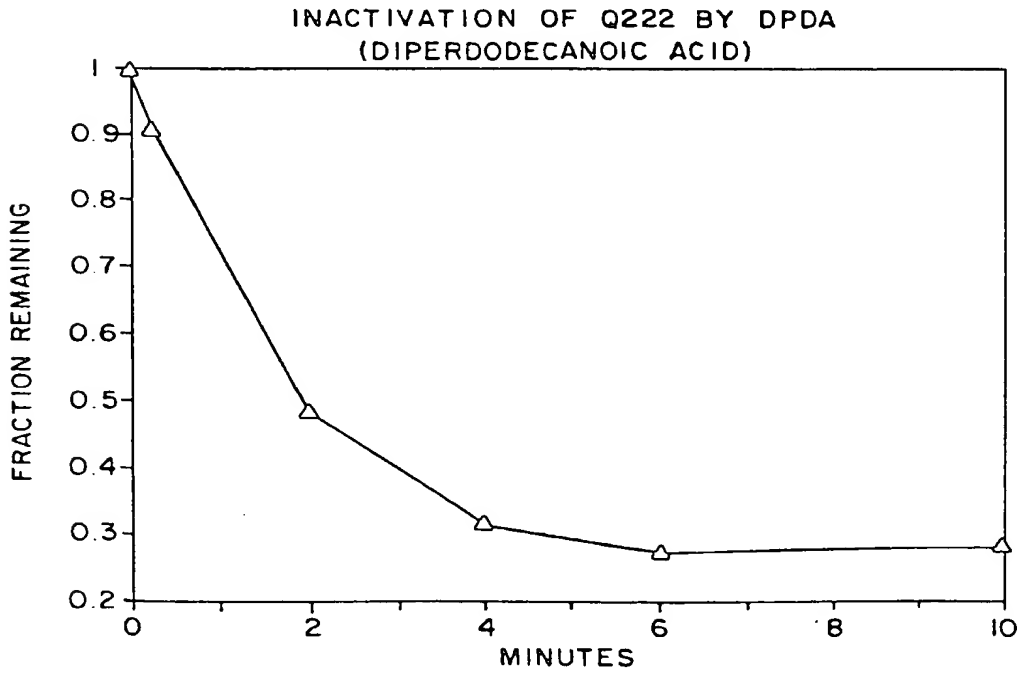

INACTIVATION OF Q222 BY DPDA
(DIPERDODECANOIC ACID)
FRACTION REMAINING
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0
2
4
6
8
10
MINUTES

FIG.—6B

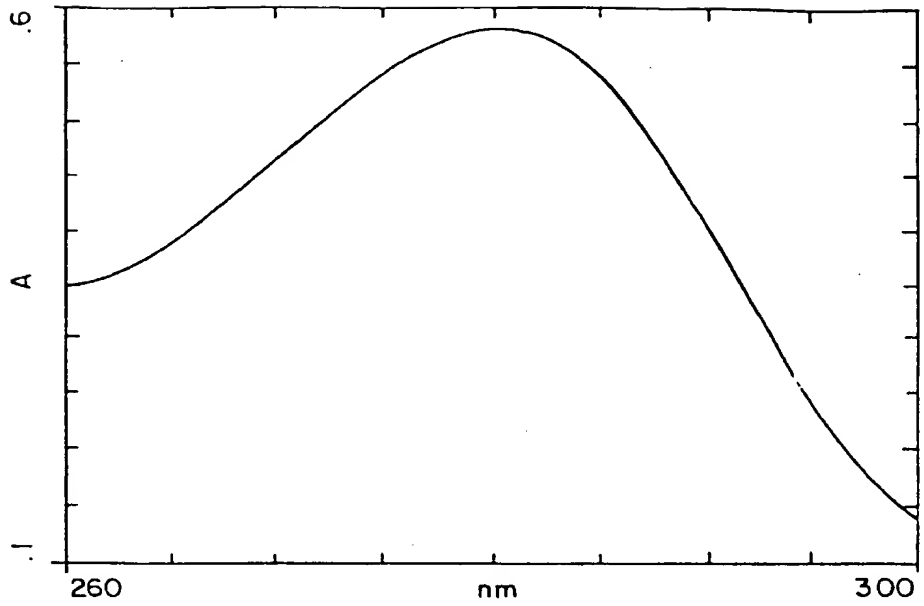


FIG. —7A

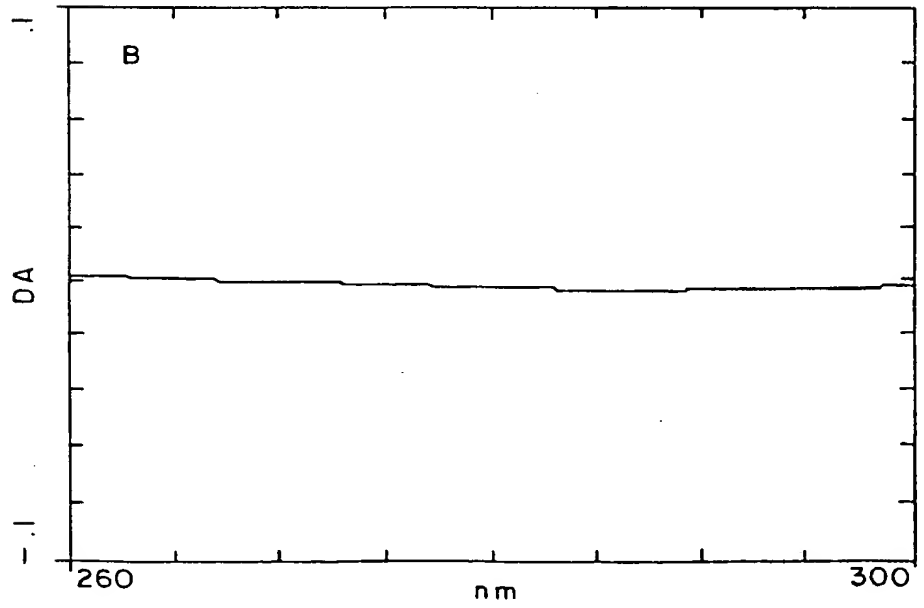


FIG. — 7B

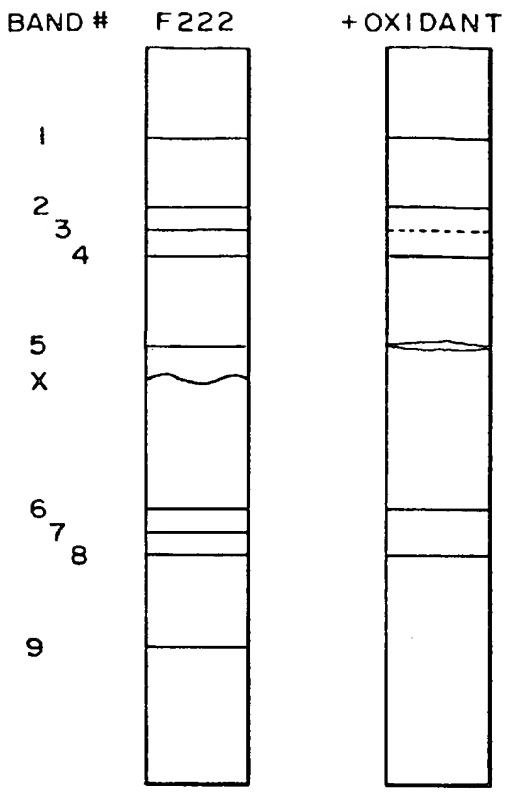

BAND #
F222
+OXIDANT
1
2
3
4
5
X
6
7
8
9

FIG. — 8

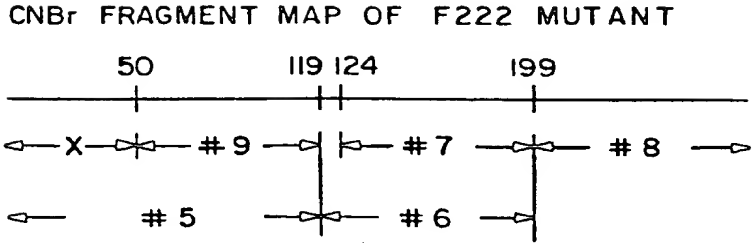

CNBr FRAGMENT MAP OF F222 MUTANT
50
119 124
199
X
#9
#7
#8
#5
#6

FIG. — 9

1. Codon number: 43 45

2. Wild type amino acid sequence: Lys-Val-Ala-Gly-Gly-Ala-Ser-Met-Val-Pro-Ser

3. Wild type DNA sequence:
```
5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTT-CCT-TCT
   TTC-CAT-CGT-CCG-CCT-CGG-TCG-TAC-CAA-GGA-AGA-5'
```

4. pΔ50:
```
        ** *                                *
5'-AAG-GCC-T----------------GC-ATG-GTA-CCT-TCT
   TTC-CGG-A----------------CG-TAC-CAT-GGA-AGA-5'
      Stu I                              Kpn I
```

5. pΔ50 cut with *Stu I/Kpn I*
```
                                           *
5'-AAG-G                                  pCT-TCT
   TTC-Cp                                CAT-GGA-AGA-5'
```

6. Cut pΔ50 ligated with cassettes:
```
                                          *
5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTA-CCT-TCT
   TCC-CAT-CGT-CCG-CCT-CGG-TCG-TAC-CAT-GGA-AGA-5'
```

7. Mutagenesis primer for pΔ50:
```
                          ***          *
5'-CT-GAT-TTA-AAG-GCC-TGC-ATG-GTA-CCT-TCT-GA
```

8. Mutants made: V45, P45, V45/P48, E46, E48, V48, C49, C50, F50

FIG.—10

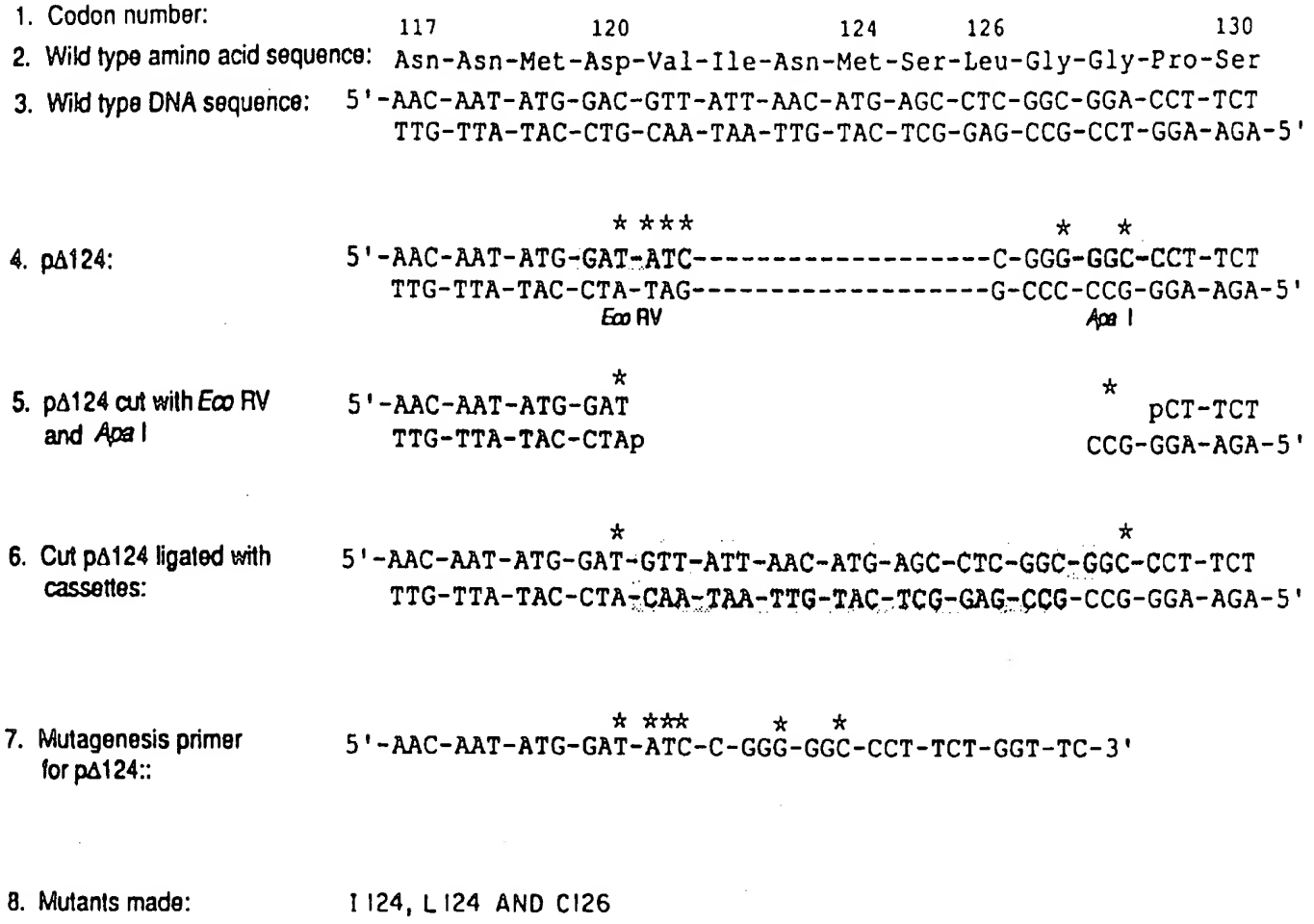


FIG.—11

EP 0 251 446 B1

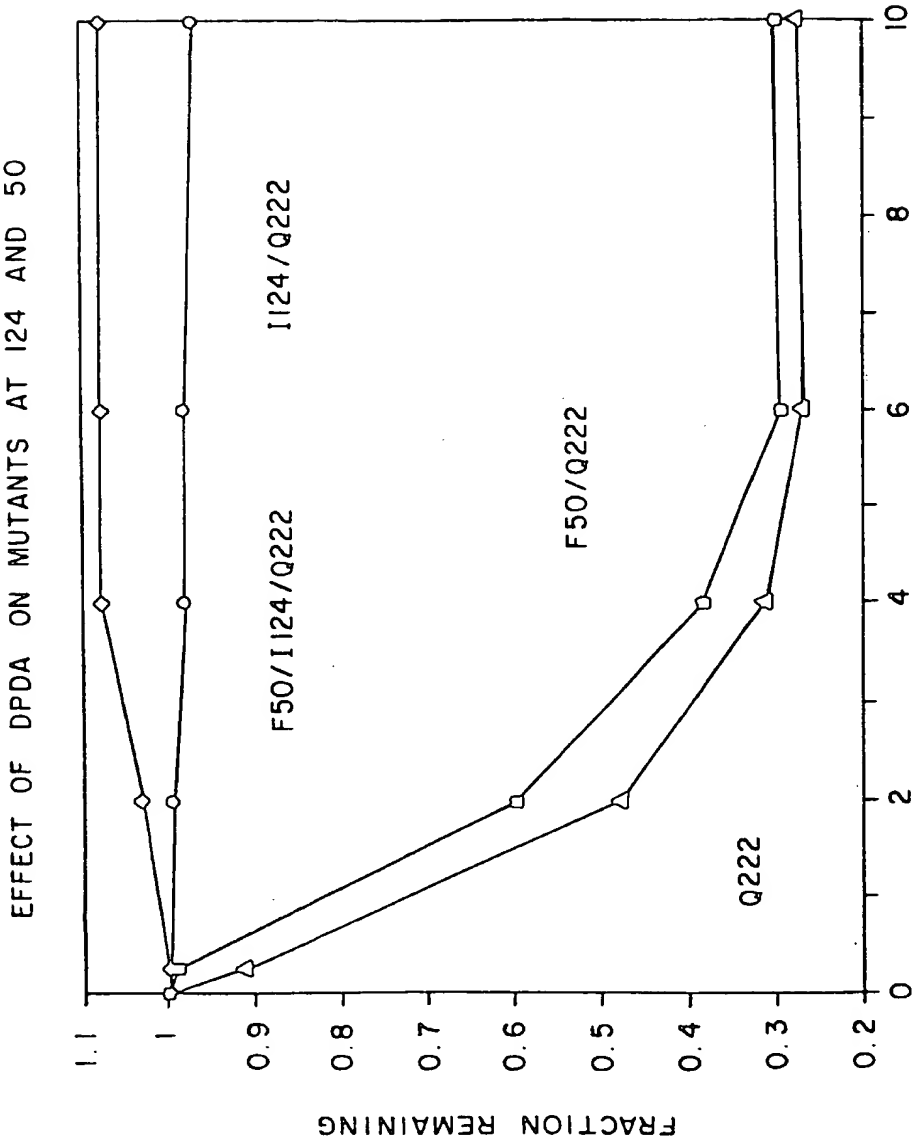

EFFECT OF DPDA ON MUTANTS AT I24 AND 50
FRACTION REMAINING
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0
2
4
6
8
10
MINUTES
I124/Q222
F50/I124/Q222
F50/Q222
Q222

FIG.-12

	Codon:	166
	Wild type amino acid sequence:	Thr Ser Gly Ser Ser Ser Thr Val Gly Tyr Pro Gly
1.	Wild type DNA sequence:	5'-ACT TCC GGC AGC TCA AGC ACA GTG GGC TAC CCT GGT-3' 3'-TGA AGG CCG TCG AGT TCG TGT CAC CCG ATG GGA CCA-5'
2.	pΔ166 DNA sequence:	5'-ACT TCC GGG AGC TCA A_ _ _ _ _ _ _ _ _ _ C CCG GGT-3' 3'-TGA AGG CCC TCG AGT T_ _ _ _ _ _ _ _ _ _ G GGC CCA-5' SacI XmaI
3.	pΔ166 cut with SacI and XmaI:	5'-ACT TCC GGG AGC T pCCG GGT-3' 3'-TGA AGG CCCp CA-5'
4.	Cut pΔ166 ligated with duplex DNA cassette pools:	5'-ACT TCC GGG AGC TCA AGC ACA GTG NNN TAC CCG GGT-3' 3'-TGA AGG CCC TCG AGT TCG TGT CAC NNN ATG GGC CCA-5'

MUTAGENESIS PRIMER 37MER

5' AA GGC ACT TCC GGG AGC TCA ACC CGG GTA AA TAC CCT 3'

FIG.—13

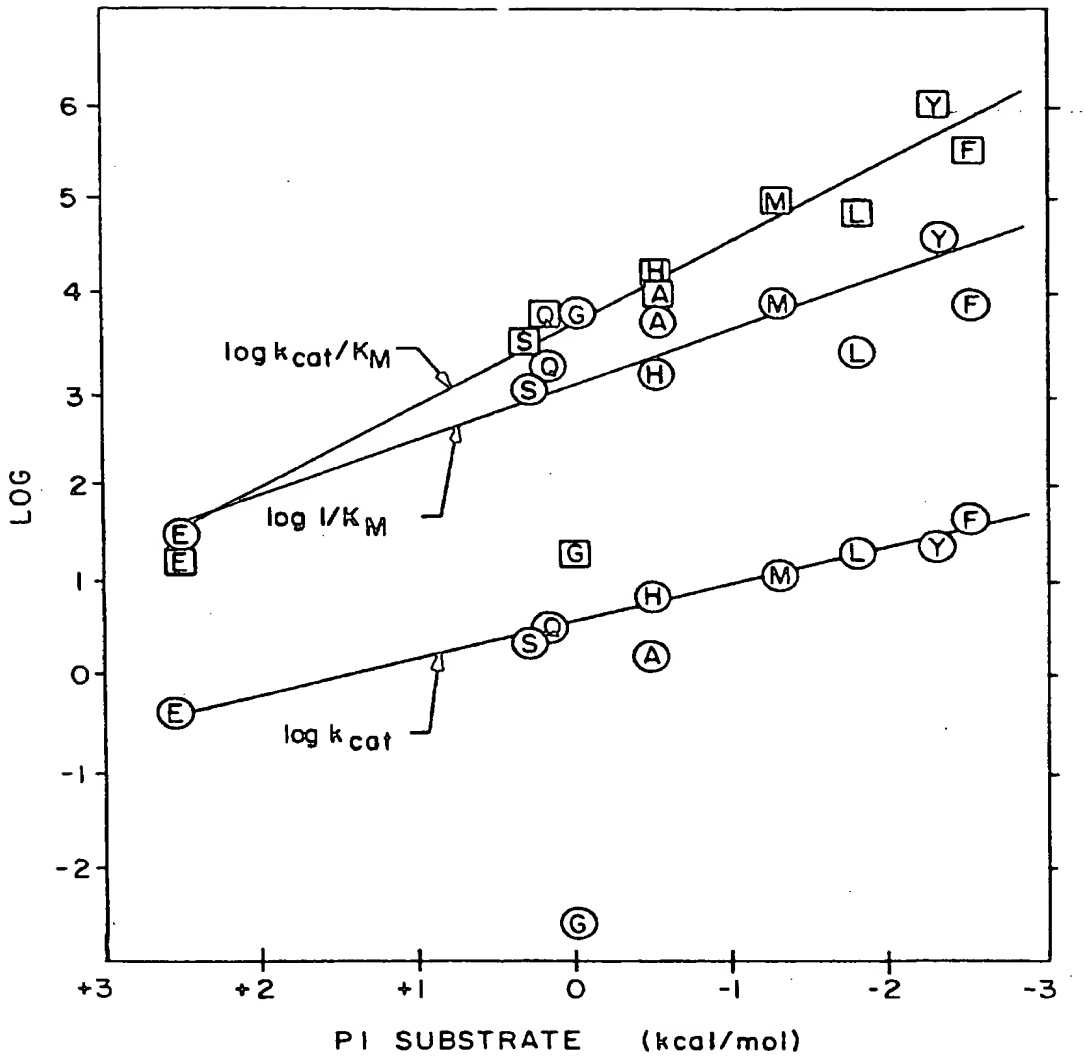

LOG
6
5
4
3
2
1
0
-1
-2
+3
+2
+1
0
-1
-2
-3
P1 SUBSTRATE (kcal/mol)
log kcat/KM
log 1/KM
log kcat

FIG. — 14

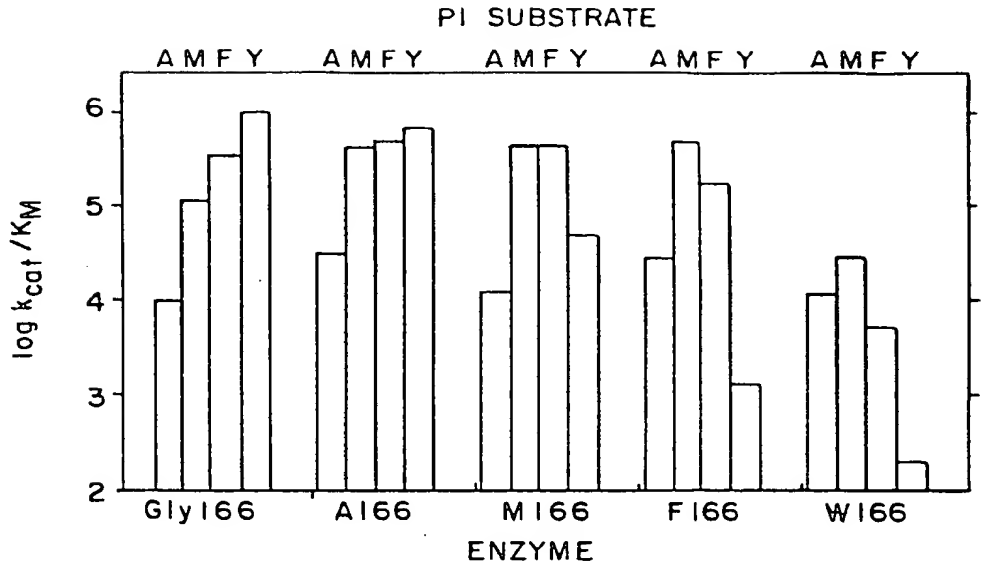


FIG.—15A

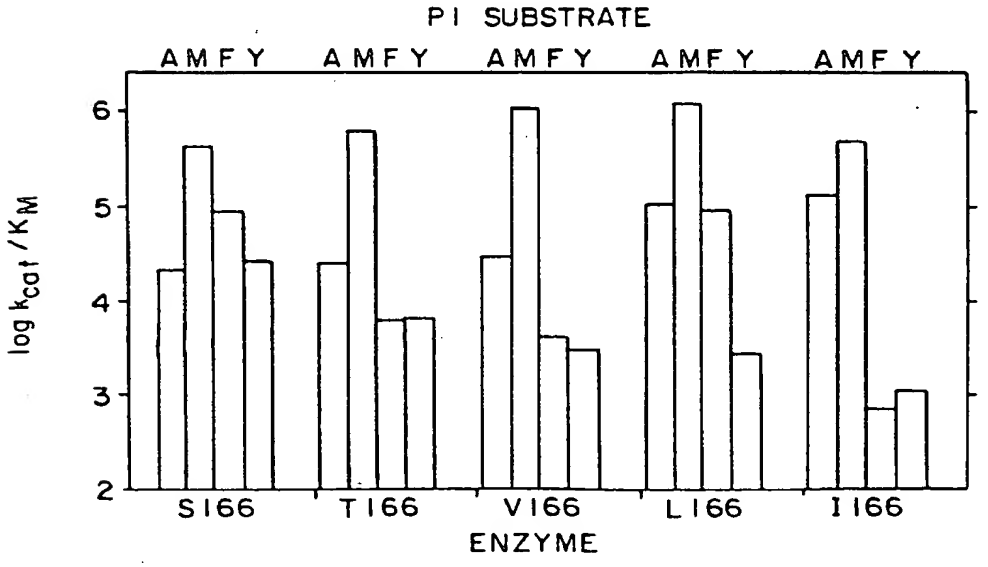


FIG.—15B

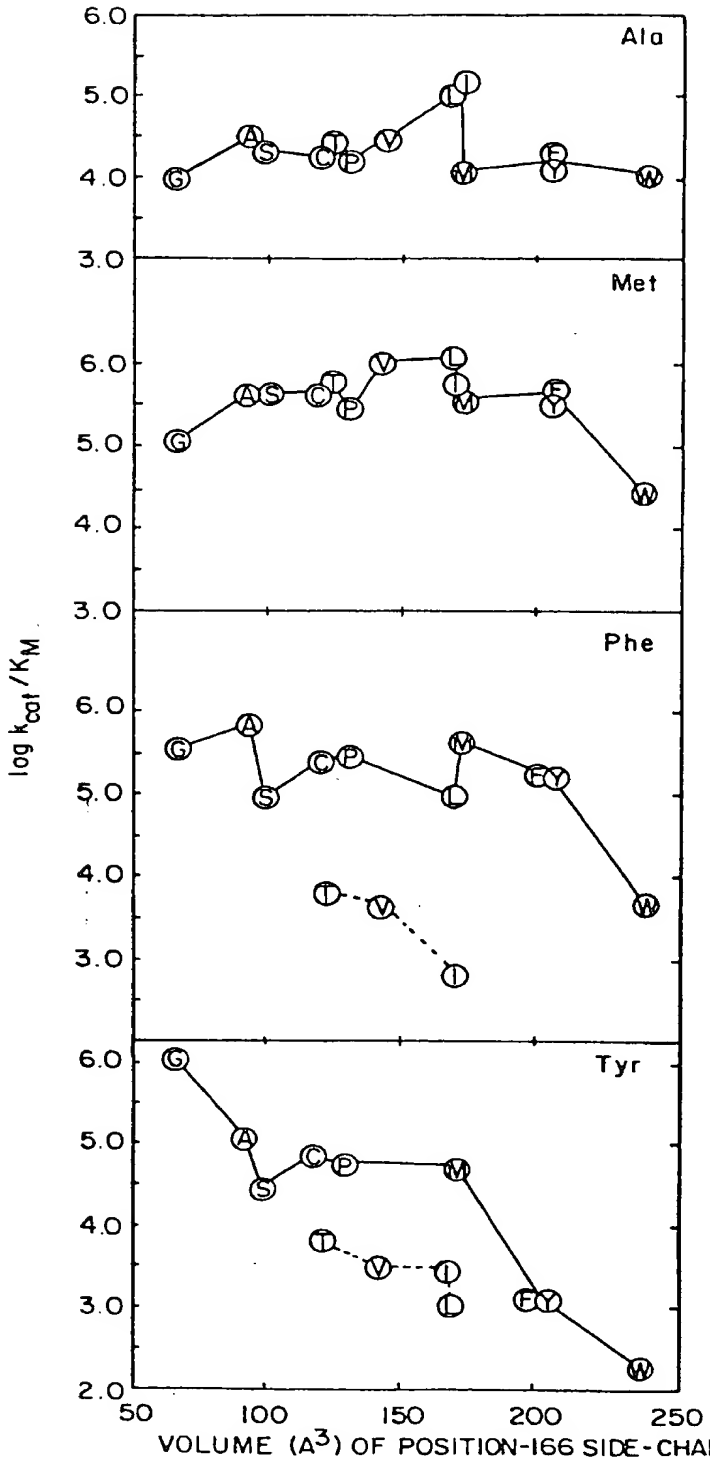

Ala
Met
Phe
Tyr
log k_{cat}/K_M
VOLUME (A^3) OF POSITION-166 SIDE-CHAIN
6.0
5.0
4.0
3.0
2.0
50
100
150
200
250

FIG.-16

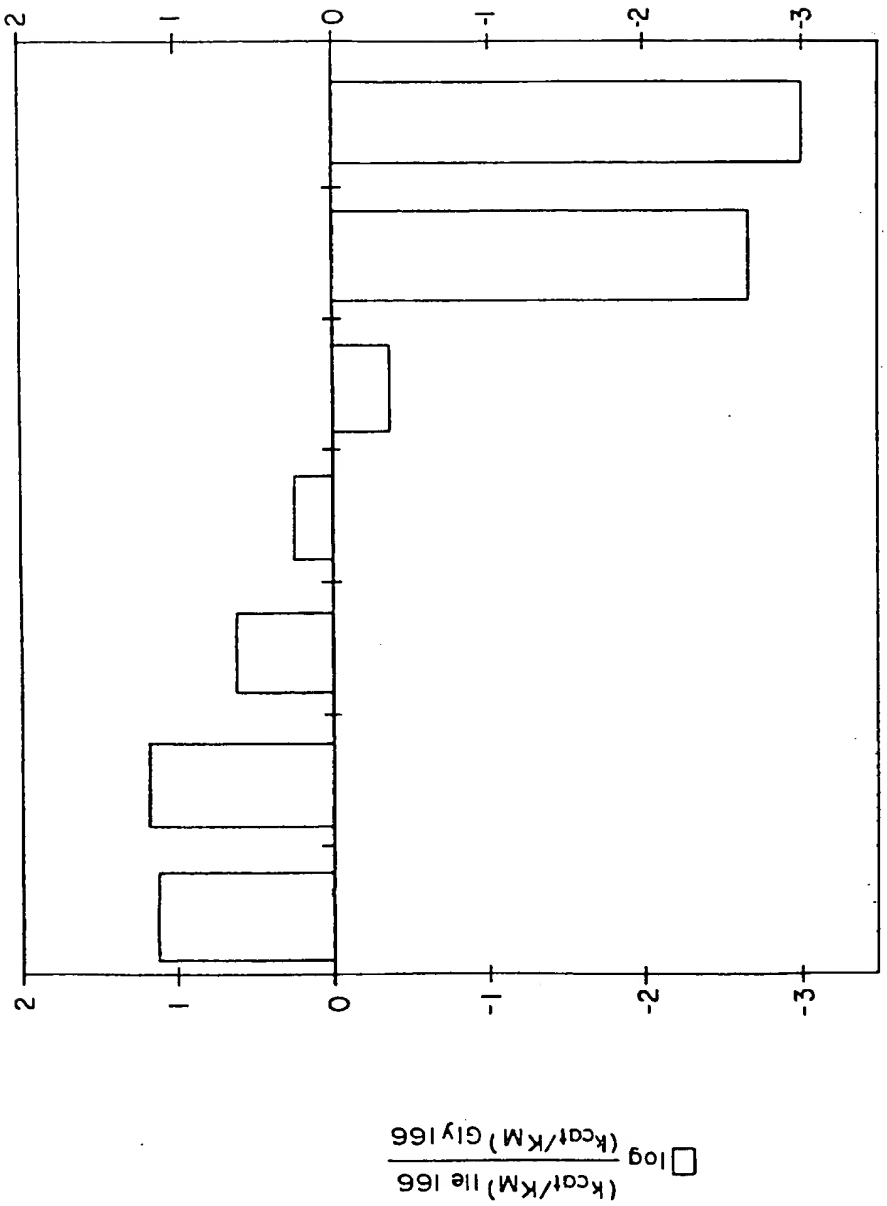

2
1
0
-1
-2
-3
Ala
Val
Met
Leu
His
Phe
Tyr
P1 SUBSTRATE
□log (kcat/KM) Ile 166 / (kcat/KM) Gly 166

FIG. - 17

GLY-169 CASSETTE MUTAGENESIS

CODON:		162 169 173
WILD TYPE AMINO ACID SEQUENCE:		SER SER THR VAL GLY TYR PRO GLY LIS TYR PRO SER
1. WILD TYPE DNA SEQUENCE	5'	TCA AGC ACA GTG GGC TAC CCT GGT AAA TAC CCT TCT 3'
	3'	AGT TCG TGT CAC CCG ATG GGA CCA TTT ATG GGA AGA 5'
2. P169 DNA SEQUENCE	5'	TCA AGC ACA GTC GGG TAC CCT-----GA TAT CCT TCT 3'
	3'	AGT TCG TGT CAC CCC ATG GGA CT ATA GGA AGA 5'
		KPNI EcoRV
3. P169 CUT WITH KPNI AND ECORV:	5'	TAC AGC ACA GTC GGG TAC PAT CCT TCT 3'
	3'	AGT TCG TGT CAC CCP TA GGA AGA 5'
4. CUT P169 LIGATED WITH	5'	TAC AGC ACA GTG GGG TAC CCT NNN AAA TAT CCT TGT 3'
OLIGONUCLEOTIDE POOLS	3'	AGT TCG TGT CAC CCC ATG GGA NNN TTT ATA GGA AGA 5'
MUTAGENESIS PRIMER FOR P169	5'	AAG CAC AGT GGG GTA CCC TGA TAT CCT TCT GTC A 3'

FIG.—18

EP 0 251 446 B1

1. Codon number: 100 104 105 108

2. Wild type amino acid sequence: Gly-Ser-Gly-Gln-Tyr-Ser-Trp-Ile-Ile-

3. Wild type DNA sequence: 5'-GGT-TCC-GGC-CAA-TAC-AGC-TGG-ATC-ATT-3'
Pvu II

4. Primer for *Hind* III insertion at 104: 5'-GGT-TCC-GGC-CAA-GCTT-AGC-TGG-ATC-ATT-3'
Hind III

5. Primers for 104 mutants: 5'---T-TCC-GCC-CAA-NNN-AGC-TGG-ATC------3'

6. Mutants made: A, M, L, S, AND HIO4

FIG.—19

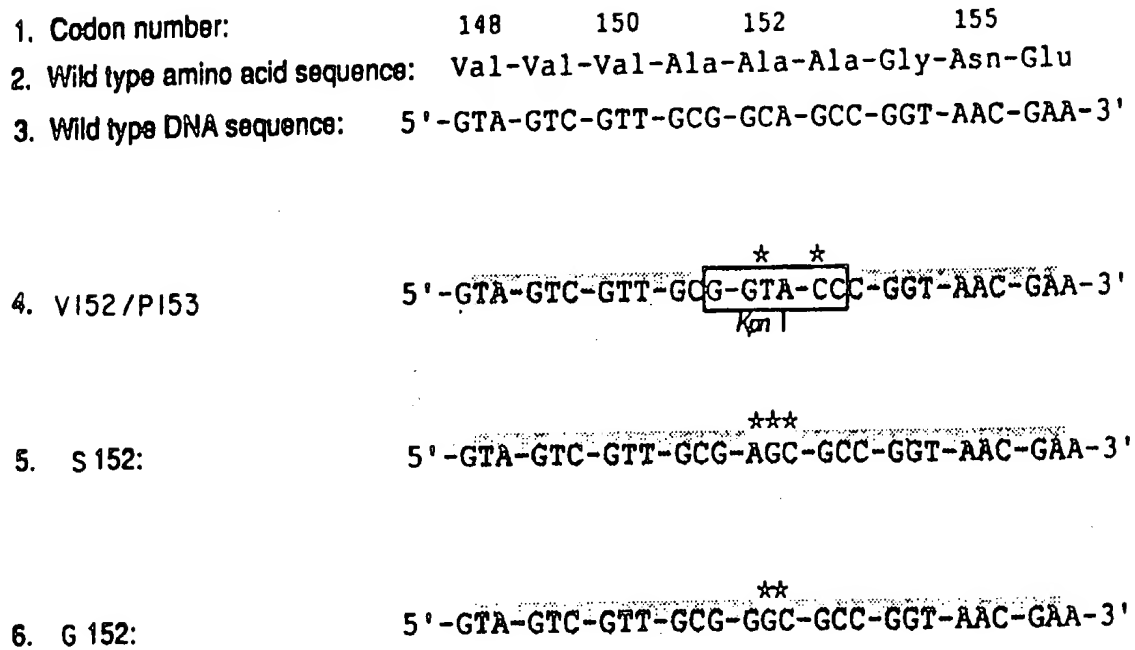

1. Codon number: 148 150 152 155
2. Wild type amino acid sequence: Val-Val-Val-Ala-Ala-Ala-Gly-Asn-Glu
3. Wild type DNA sequence: 5'-GTA-GTC-GTT-GCG-GCA-GCC-GGT-AAC-GAA-3'
4. V152/P153 5'-GTA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3'
Kpn I
5. S 152: 5'-GTA-GTC-GTT-GCG-AGC-GCC-GGT-AAC-GAA-3'
6. G 152: 5'-GTA-GTC-GTT-GCG-GGC-GCC-GGT-AAC-GAA-3'

FIG.—20

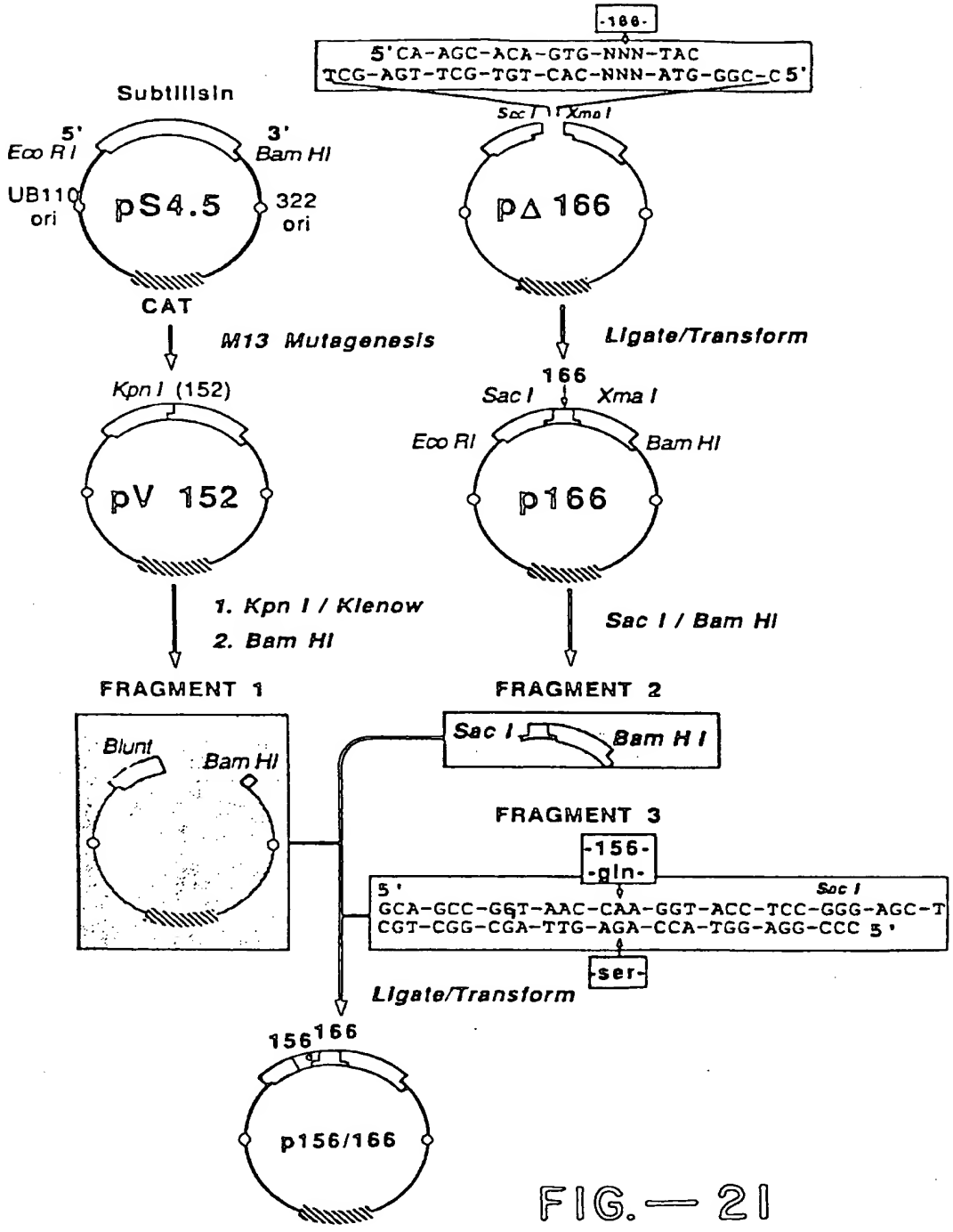

-166-
5'CA-AGC-ACA-GTG-NNN-TAC
TCG-AGT-TCG-TGT-CAC-NNN-ATG-GGC-C 5'
Subtilisin
5'
Eco RI
3'
Bam HI
UB110 ori
pS4.5
322 ori
CAT
Sac I
Xma I
pΔ 166
M13 Mutagenesis
Ligate/Transform
Kpn I (152)
pV 152
166
Sac I
Xma I
Eco RI
Bam HI
p166
1. Kpn I / Klenow
2. Bam HI
Sac I / Bam HI
FRAGMENT 1
FRAGMENT 2
Blunt
Bam HI
Sac I
Bam H I
FRAGMENT 3
-156-
-gln-
5'
Sac I
GCA-GCC-GGT-AAC-CAA-GGT-ACC-TCC-GGG-AGC-T
CGT-CGG-CGA-TTG-AGA-CCA-TGG-AGG-CCC 5'
-ser-
Ligate/Transform
156 166
p156/166

FIG.— 21

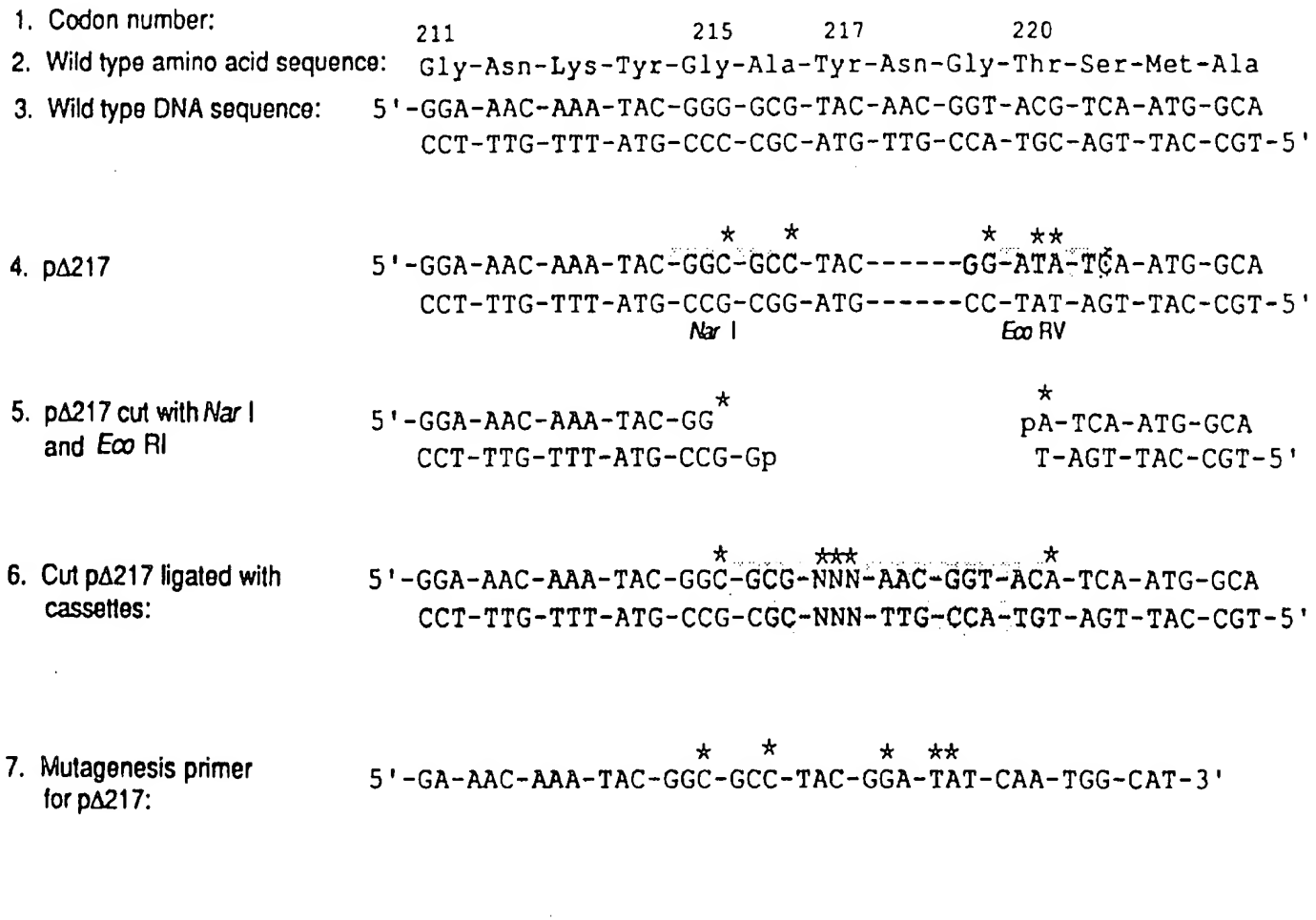

1. Codon number:
211
215
217
220
2. Wild type amino acid sequence:
Gly-Asn-Lys-Tyr-Gly-Ala-Tyr-Asn-Gly-Thr-Ser-Met-Ala
3. Wild type DNA sequence:
5'-GGA-AAC-AAA-TAC-GGG-GCG-TAC-AAC-GGT-ACG-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCC-CGC-ATG-TTG-CCA-TGC-AGT-TAC-CGT-5'
4. pΔ217
5'-GGA-AAC-AAA-TAC-GGC-GCC-TAC------GG-ATA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGG-ATG------CC-TAT-AGT-TAC-CGT-5'
Nar I
Eco RV
5. pΔ217 cut with Nar I and Eco RI
5'-GGA-AAC-AAA-TAC-GG
pA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-Gp
T-AGT-TAC-CGT-5'
6. Cut pΔ217 ligated with cassettes:
5'-GGA-AAC-AAA-TAC-GGC-GCG-NNN-AAC-GGT-ACA-TCA-ATG-GCA
CCT-TTG-TTT-ATG-CCG-CGC-NNN-TTG-CCA-TGT-AGT-TAC-CGT-5'
7. Mutagenesis primer for pΔ217:
5'-GA-AAC-AAA-TAC-GGC-GCC-TAC-GGA-TAT-CAA-TGG-CAT-3'
8. Mutants made:
All 19 at 217

FIG.—22

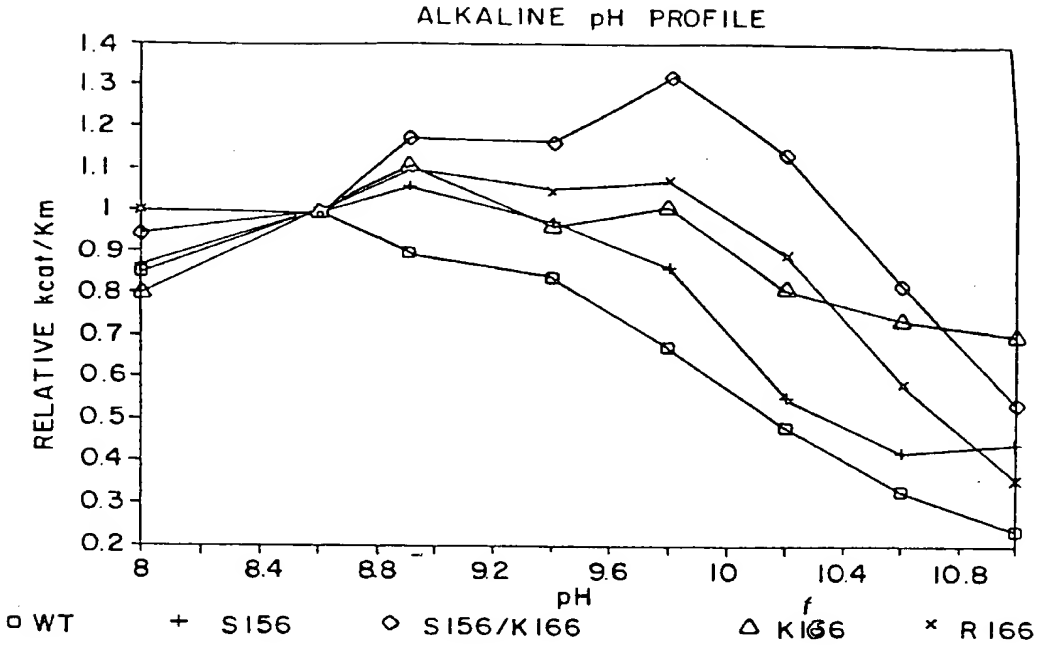


FIG. — 23A

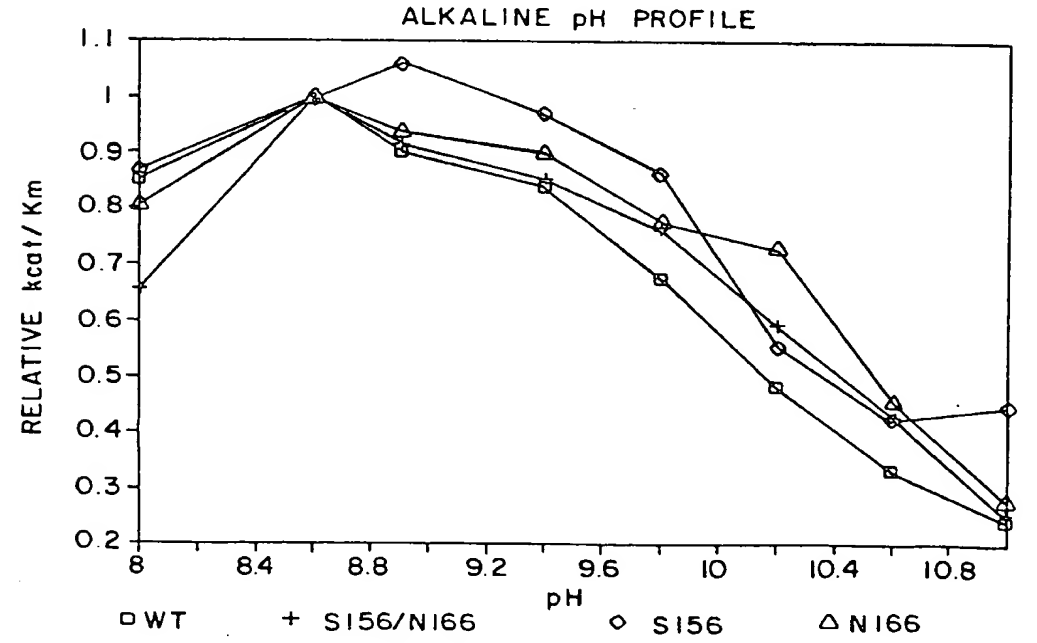


FIG. — 23B

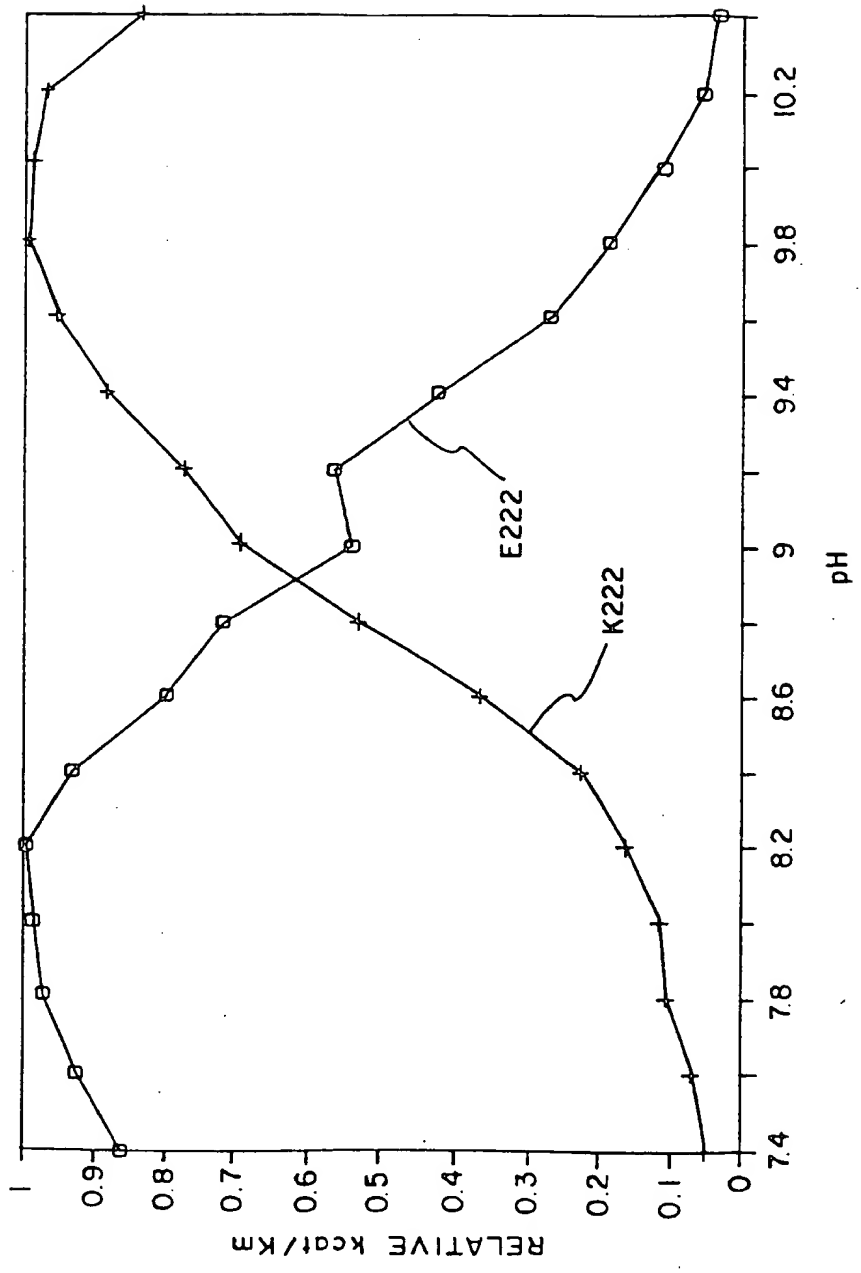

RELATIVE kcat/km
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
7.4
7.8
8.2
8.6
9
9.4
9.8
10.2
pH
E222
K222

FIG. - 24

1. Codon number: 91 95 100

2. Wild type amino acid sequence: Tyr-Ala-Val-Lys-Val-Leu-Gly-Ala-Asp-Gly-Ser

3. Wild type DNA sequence:
```
5'-TAC-GCT-GTA-AAA-GTT-CTC-GGT-GCT-GAC-GGT-TCC
   ATG-CGA-CAT-TTT-CAA-GAG-CCA-CGA-CTG-CCA-AGG-5'
```

4. pΔ95:
```
        * *                 *   *
5'-TAC-GCG-T----------CTC-GCT-GCA-GAC-GGT-TCC
   ATG-CGC-A----------GAG-CGA-CGT-CTG-CCA-AGG-5'
    Mlu I                  Pst I
```

5. pΔ95 cut with *Mlu* I *and Pst* I
```
5'-TA   *                   * pGAC-GGT-TCC
   ATG-CGCp                 A-CGT-CTG-CCA-AGG-5'
```

6. Cut pΔ95 ligated with cassettes:
```
        *                   *
5'-TAC-GCG-GTA-AAA-GTT-CTC-GGT-GCA-GAC-GGT-TCC
   ATG-CGC-CAT-TTT-CAA-GAG-CCA-CGT-CTG-CCA-AGG-5'
```

7. Mutagenesis primer for pΔ95:
```
                    * *     *   *
5'-CA-TCA-CTT-TAC-GCG-T-CTC-GCT-GCA-GAC-GGT-TCC
```

8. Mutants made: C94, C95, D96

FIG.—25

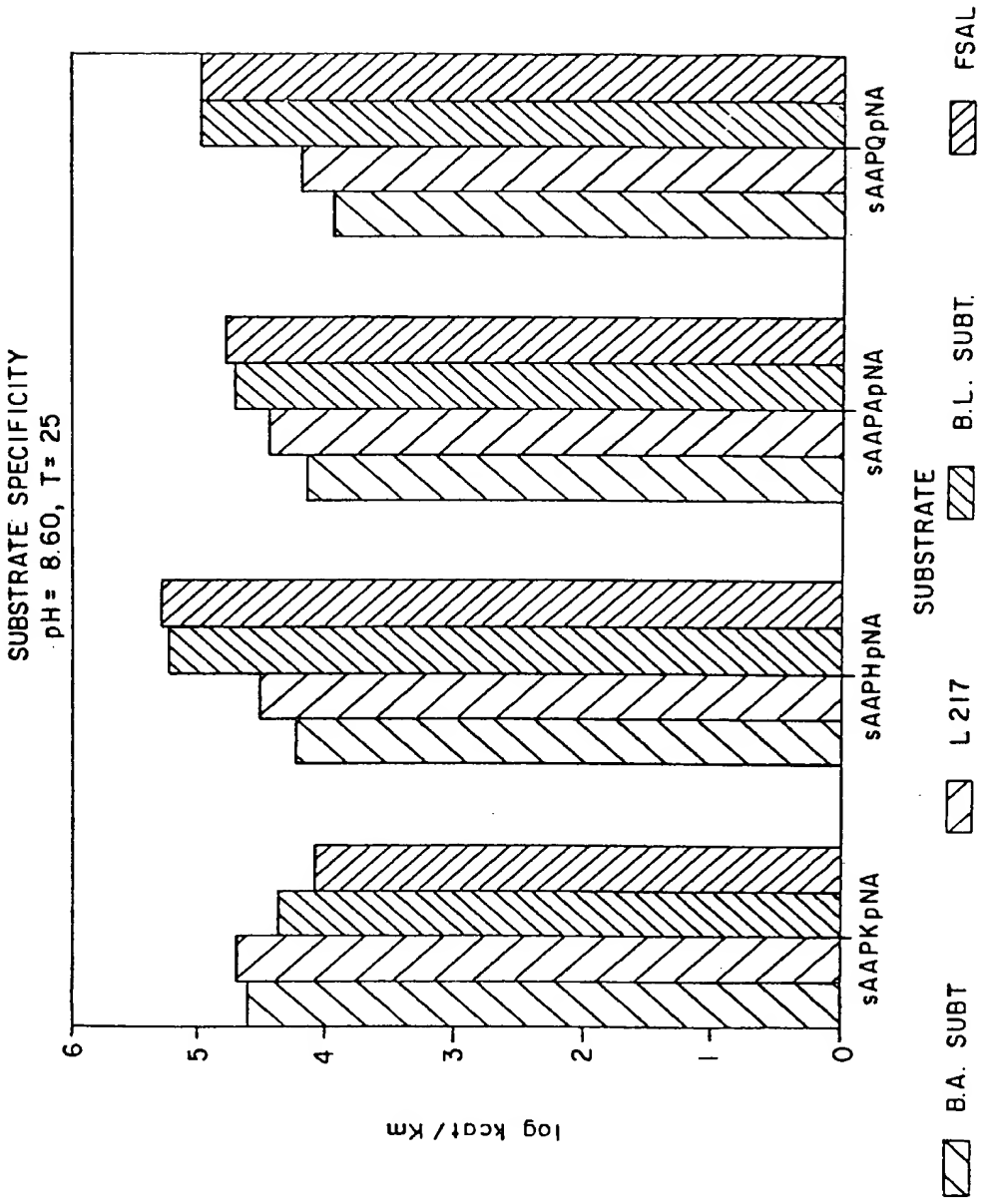

SUBSTRATE SPECIFICITY
pH= 8.60, T= 25
log kcat/ Km
0
1
2
3
4
5
6
sAAPKpNA
sAAPHpNA
sAAPApNA
sAAPQpNA
SUBSTRATE
B.A. SUBT
L 217
B.L. SUBT.
FSAL

FIG.-26

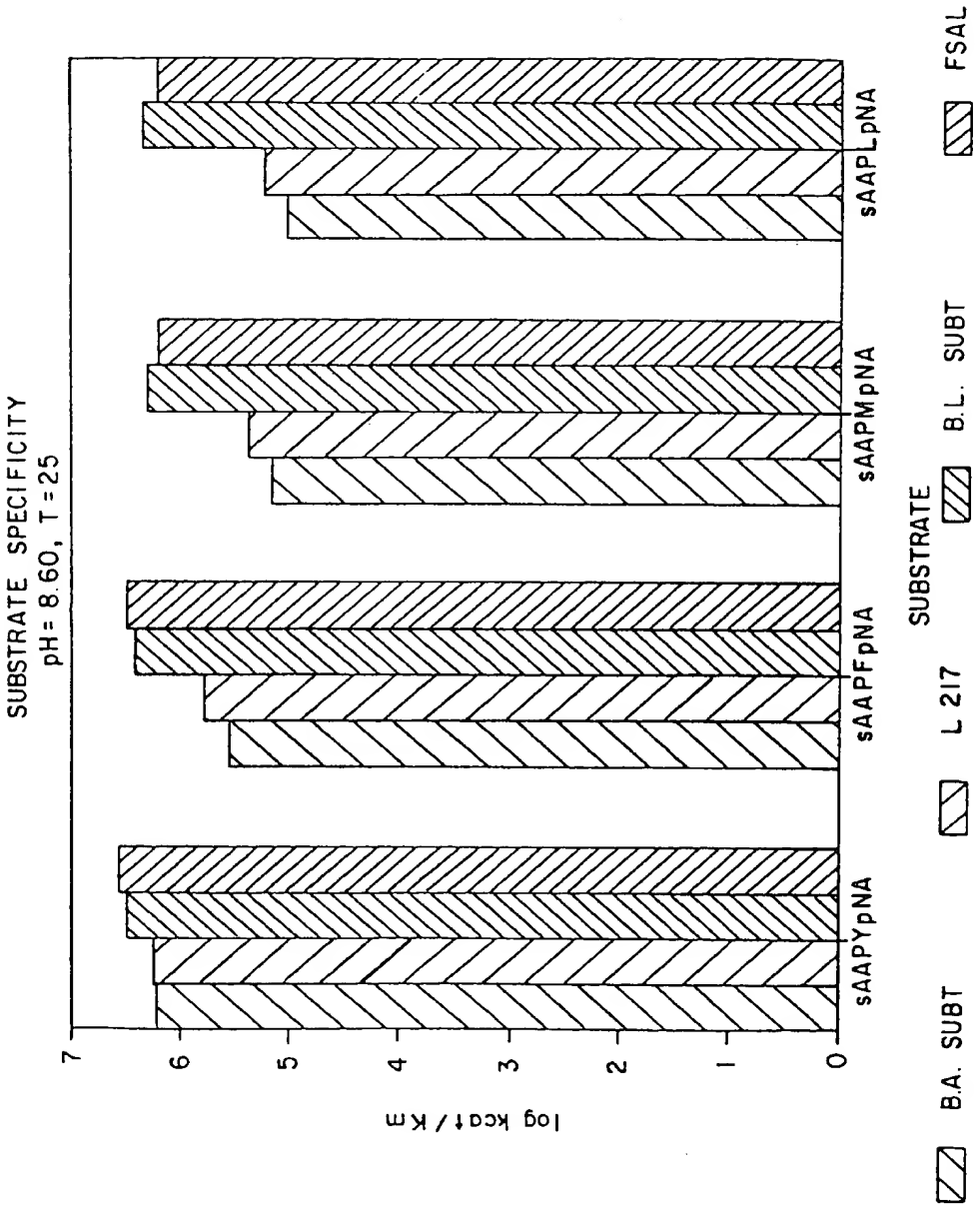

SUBSTRATE SPECIFICITY
pH= 8.60, T = 25
log kcat / Km
0
1
2
3
4
5
6
7
sAAPYpNA
sAAPFpNA
sAAPMpNA
sAAPLpNA
SUBSTRATE
B.A. SUBT
L 217
B.L. SUBT
FSAL

FIG.-27

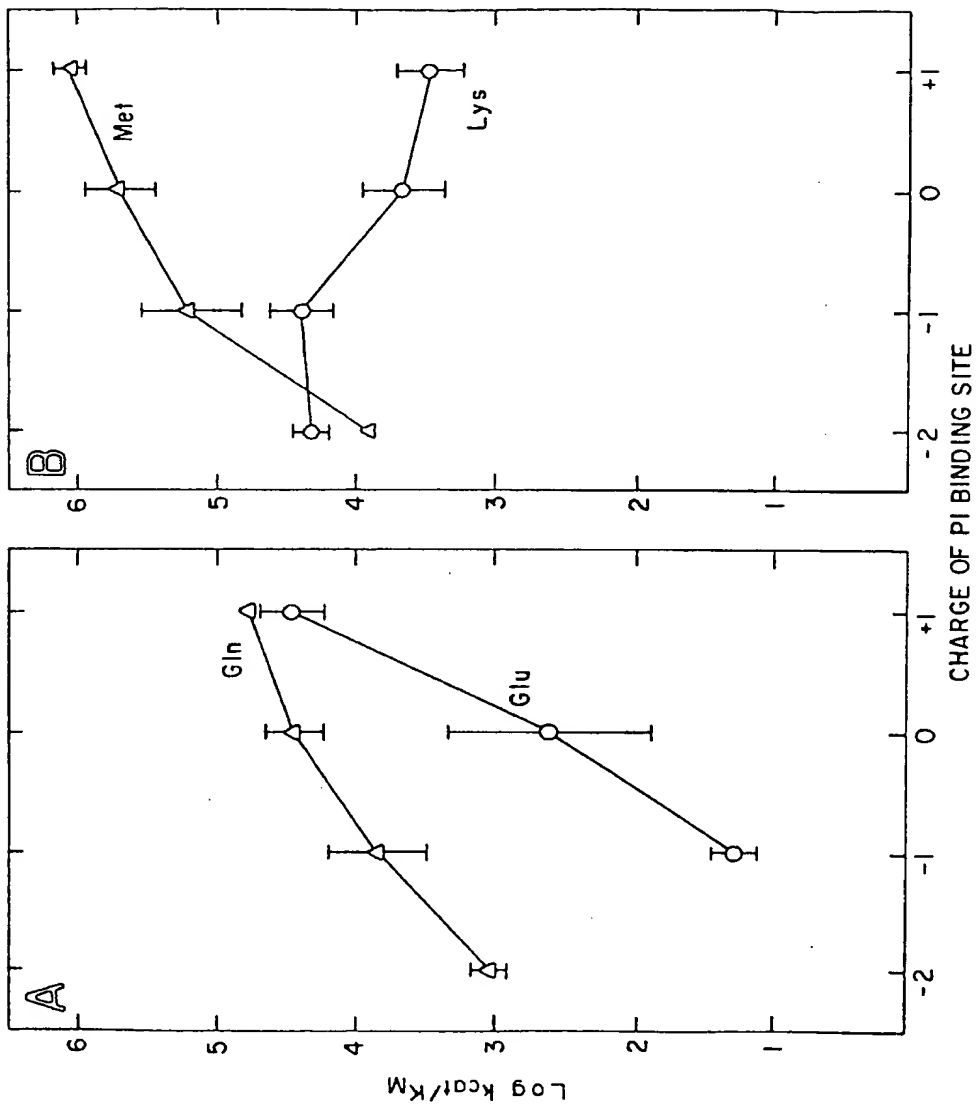

A
B
Gln
Glu
Met
Lys
Log kcat/KM
CHARGE OF PI BINDING SITE
-2
-1
0
+1
1
2
3
4
5
6

FIG. -28

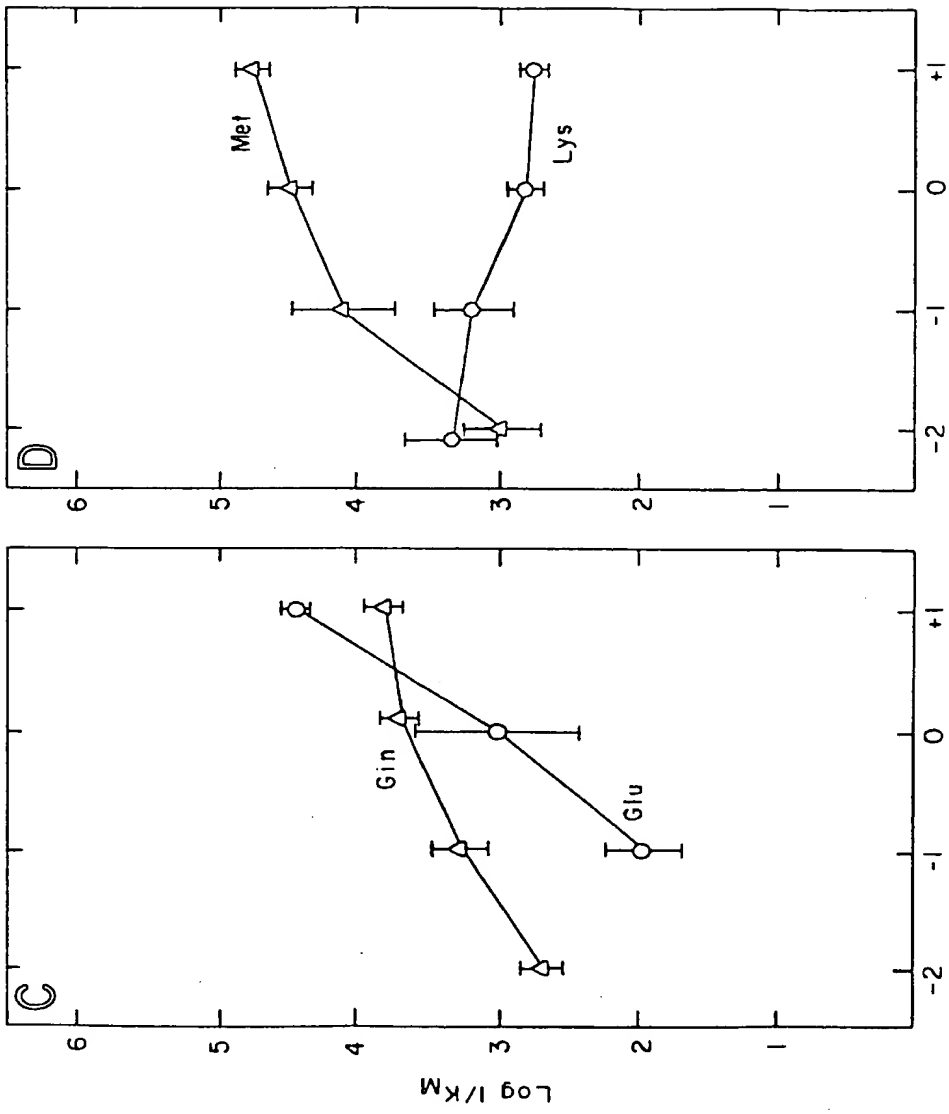

Log 1/K_M
C
D
Gln
Glu
Met
Lys
-2
-1
0
+1
1
2
3
4
5
6
CHARGE OF P1 BINDING SITE

FIG. -28

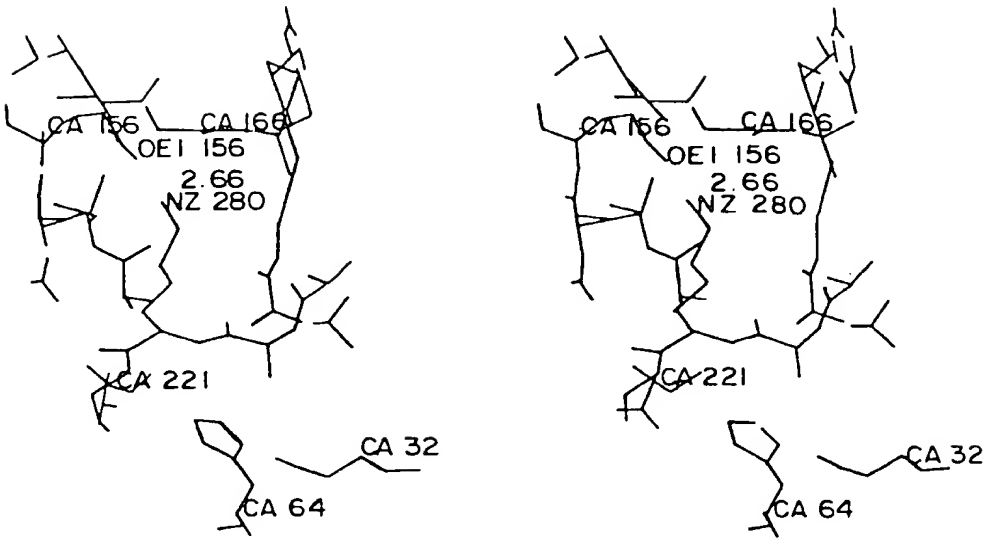


FIG. — 29A

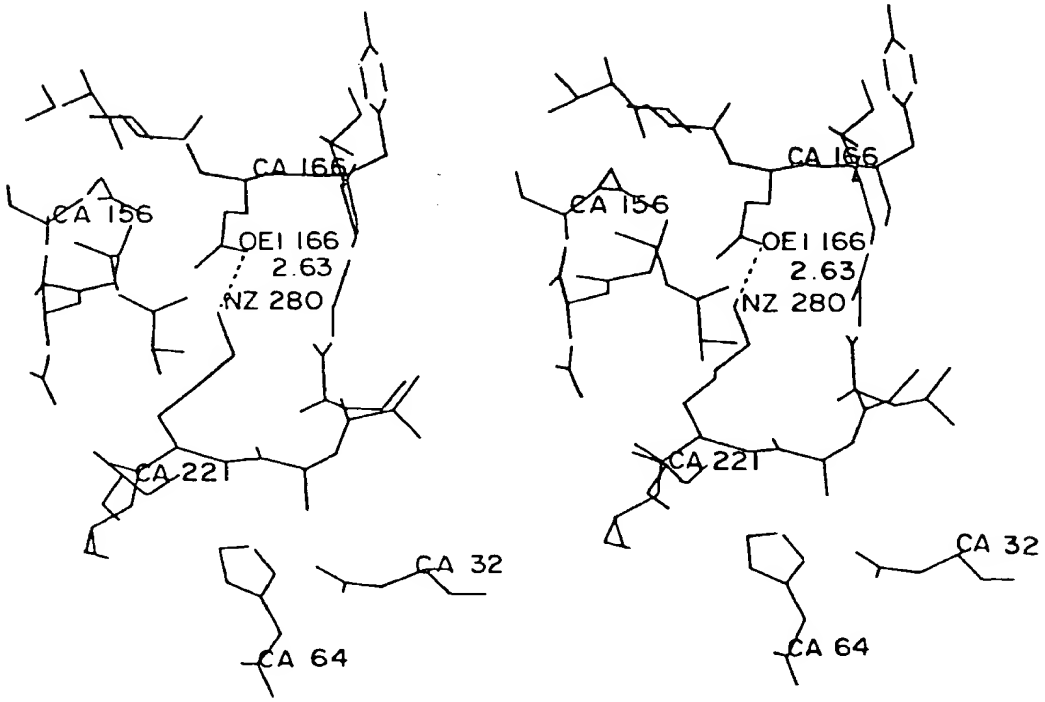


FIG. — 29B

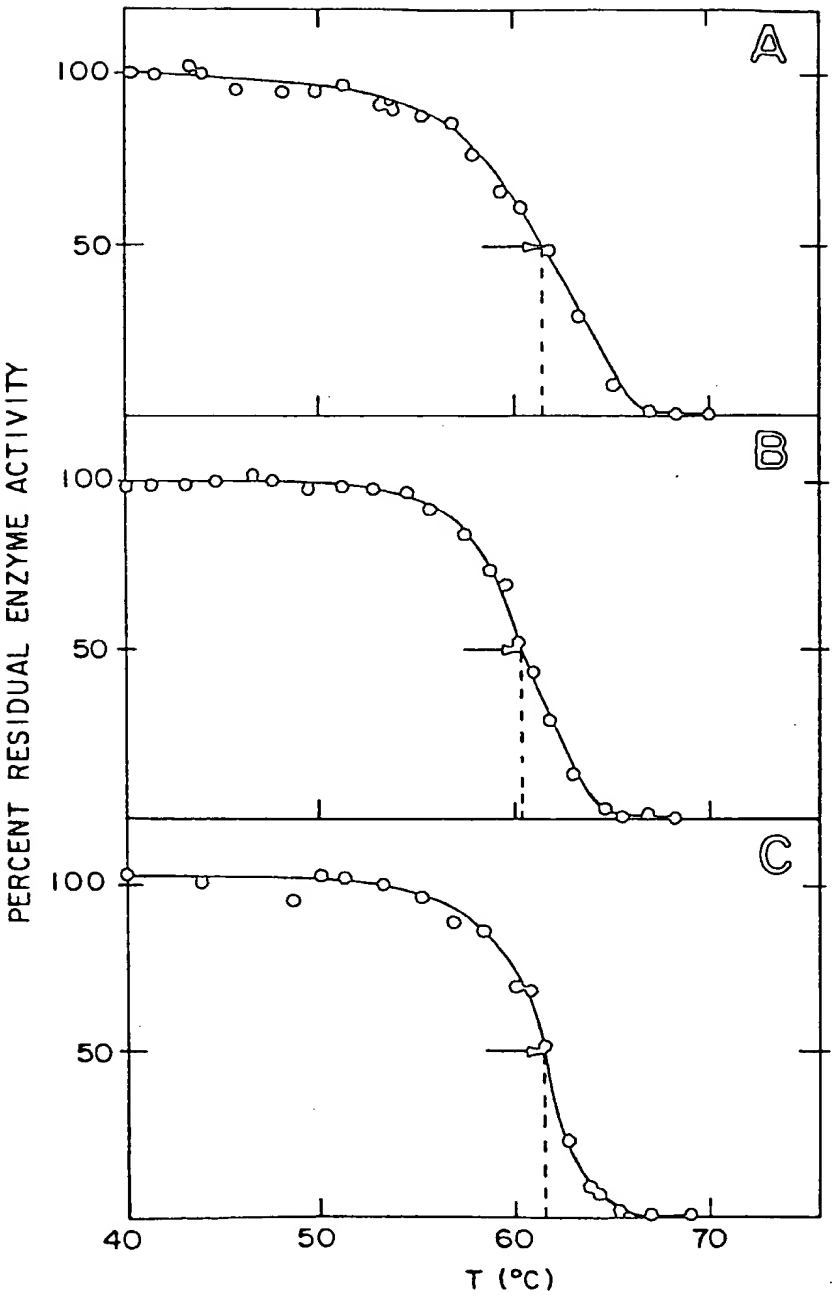

A
B
C
100
50
100
50
100
50
PERCENT RESIDUAL ENZYME ACTIVITY
40
50
60
70
T (°C)

FIG.—30

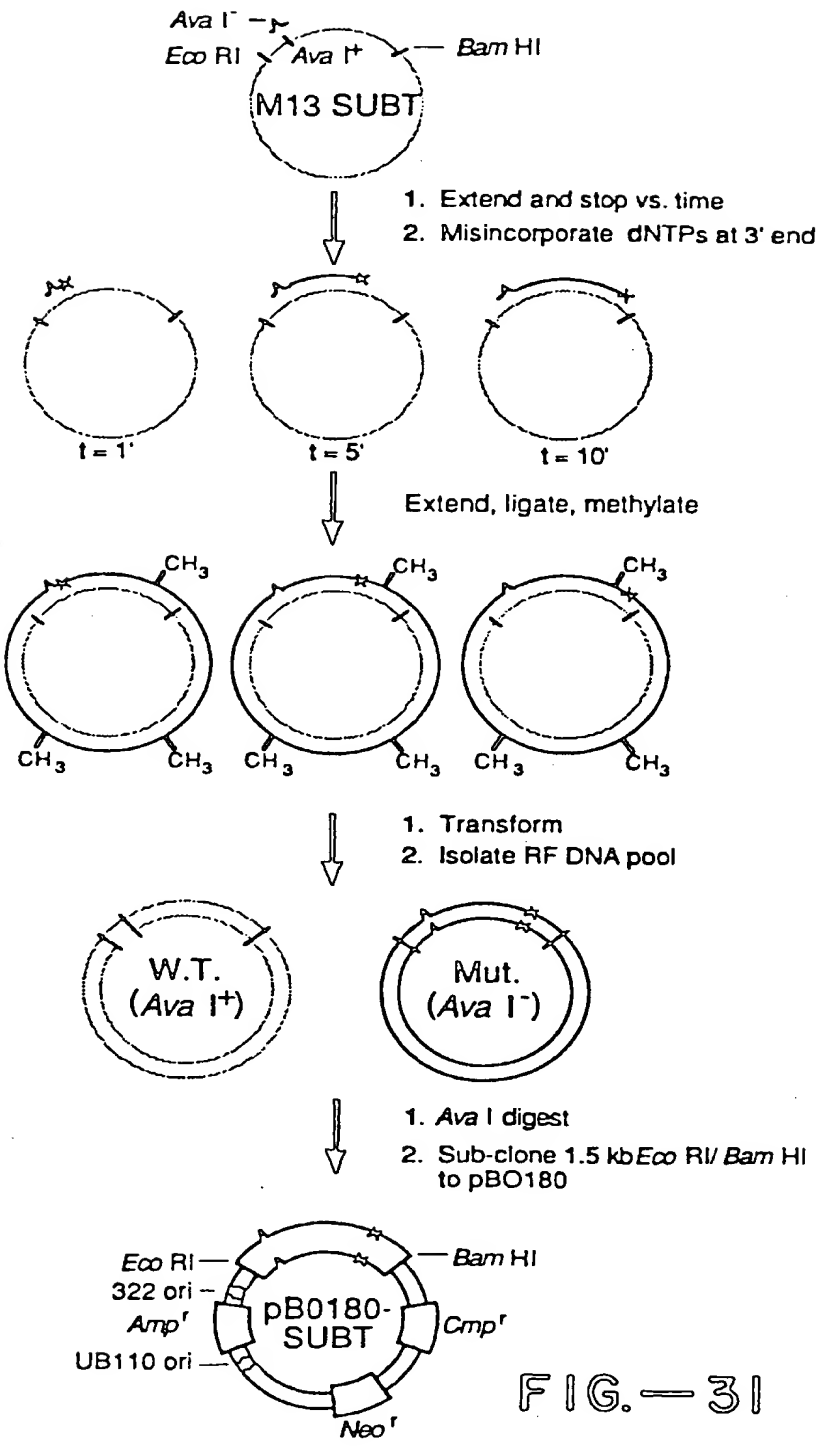

Ava I⁻
Eco RI
Ava I⁺
Bam HI
M13 SUBT
1. Extend and stop vs. time
2. Misincorporate dNTPs at 3' end
t = 1'
t = 5'
t = 10'
Extend, ligate, methylate
CH3
1. Transform
2. Isolate RF DNA pool
W.T.
(Ava I⁺)
Mut.
(Ava I⁻)
1. Ava I digest
2. Sub-clone 1.5 kb Eco RI/ Bam HI
to pBO180
Eco RI
Bam HI
322 ori
Amp^r
pB0180-
SUBT
Cmp^r
UB110 ori
Neo^r

FIG.—31

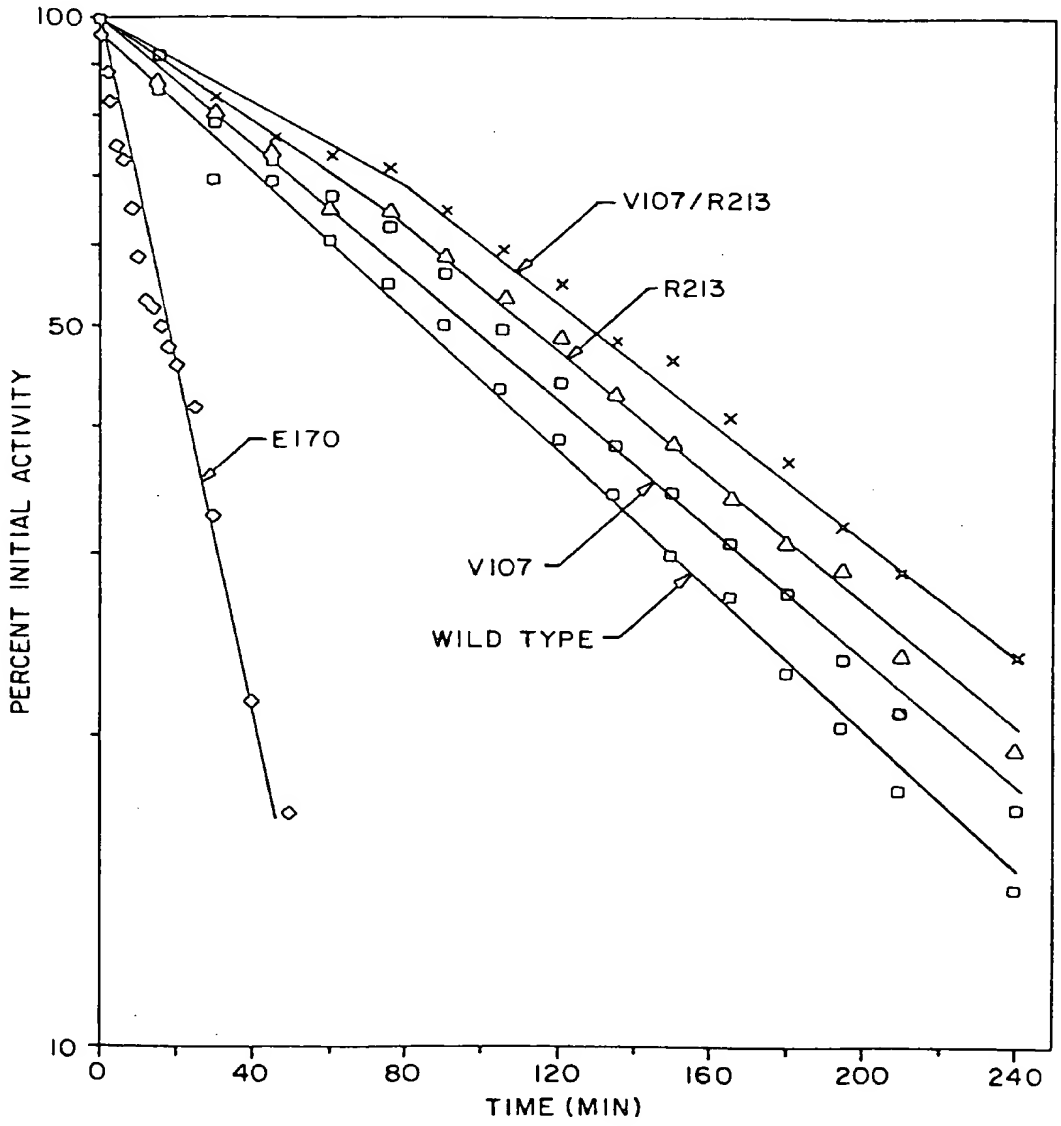

100
50
10
PERCENT INITIAL ACTIVITY
0
40
80
120
160
200
240
TIME (MIN)
V107/R213
R213
E170
V107
WILD TYPE

FIG.-32

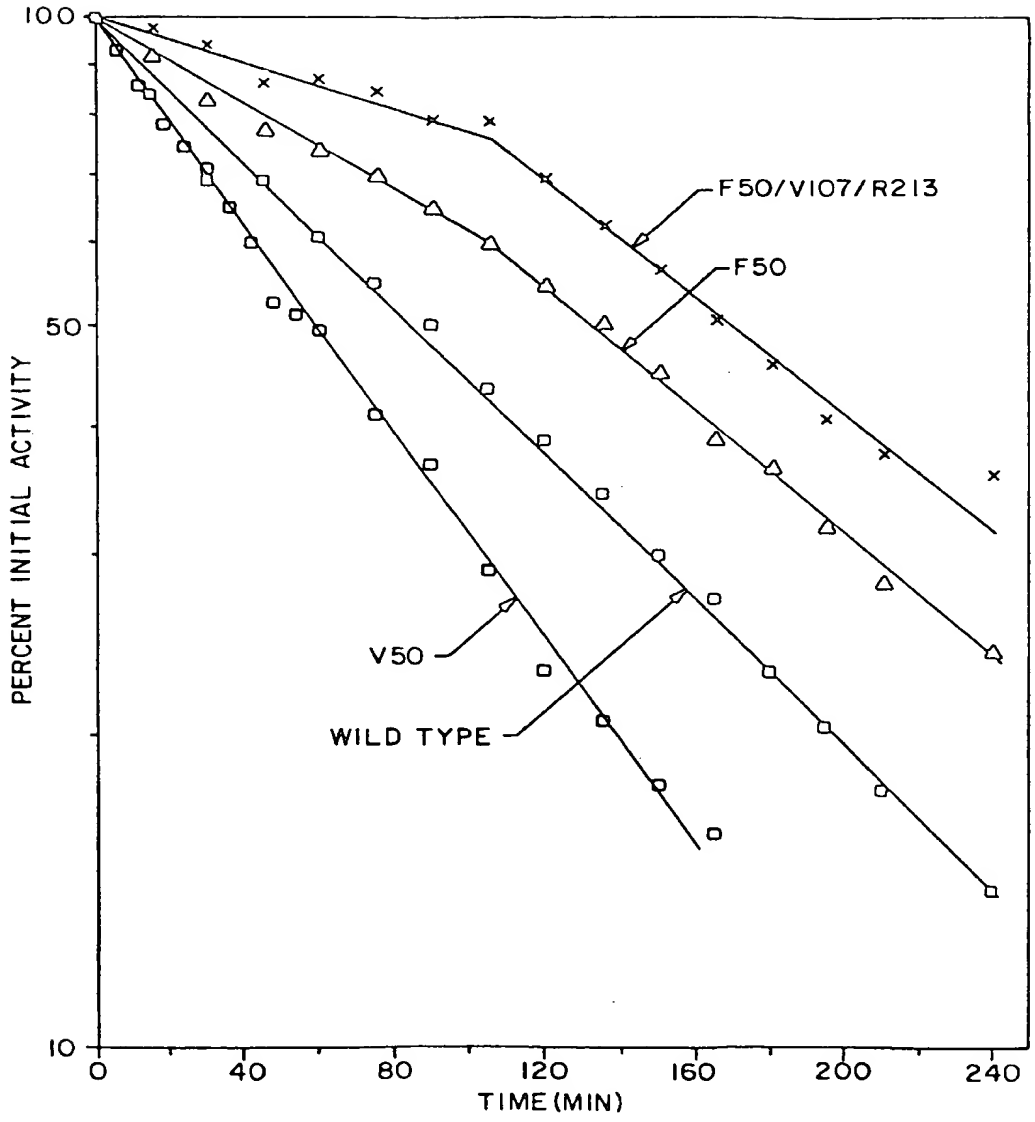

100
50
10
PERCENT INITIAL ACTIVITY
0
40
80
120
160
200
240
TIME (MIN)
F50/V107/R213
F50
V50
WILD TYPE

FIG.—33

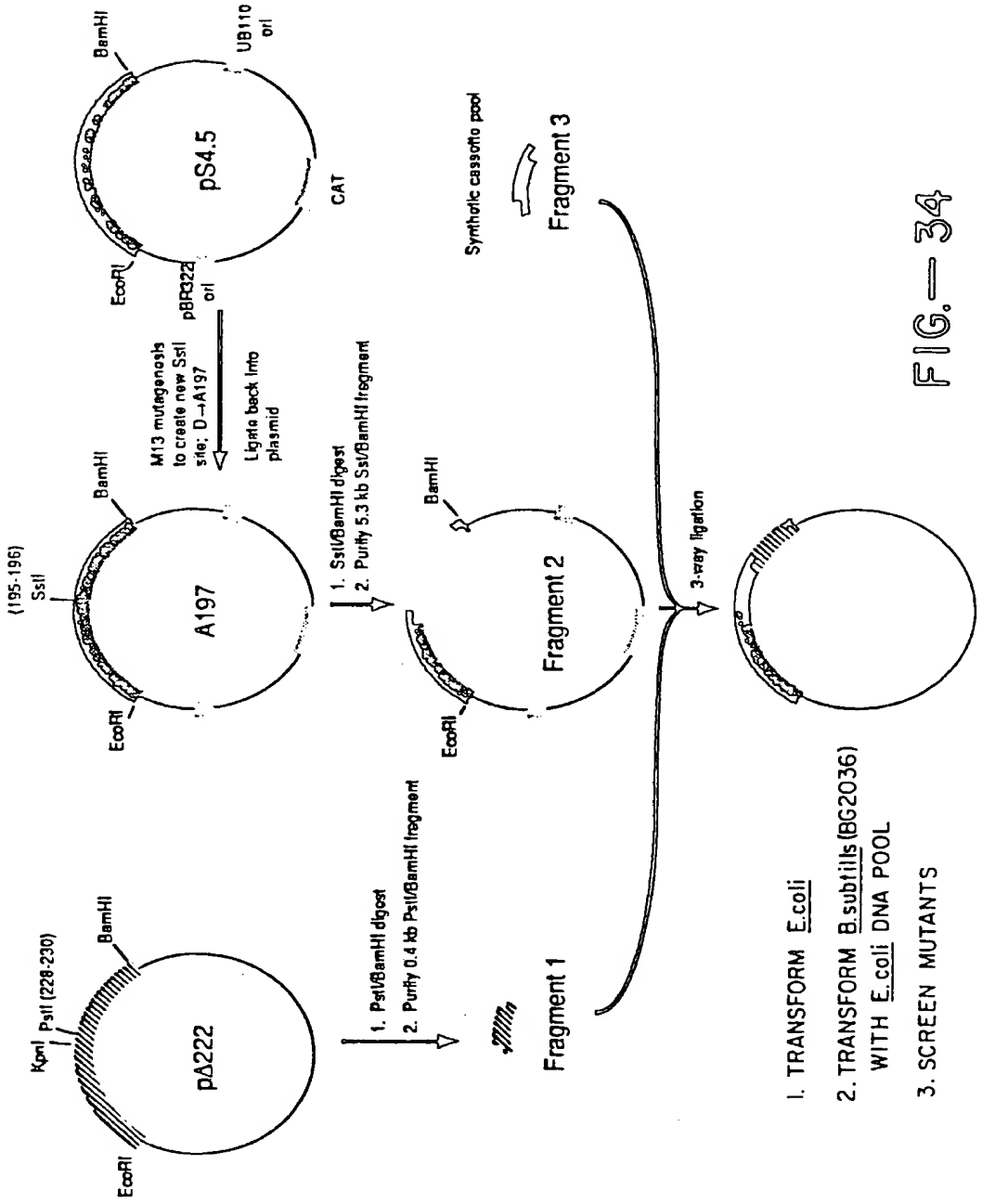

pS4.5
BamHI
UB110 ori
CAT
EcoRI
pBR322 ori
M13 mutagenesis to create new SstI site; D→A197
Ligate back into plasmid
A197
(195-196) SstI
BamHI
EcoRI
1. SstI/BamHI digest
2. Purify 5.3 kb SstI/BamHI fragment
Fragment 2
BamHI
EcoRI
Synthetic cassette pool
Fragment 3
1. PstI/BamHI digest
2. Purify 0.4 kb PstI/BamHI fragment
Fragment 1
3-way ligation
pΔ222
KpnI
PstI (228-230)
BamHI
EcoRI
1. TRANSFORM E.coli
2. TRANSFORM B.subtilis (BG2036) WITH E.coli DNA POOL
3. SCREEN MUTANTS

FIG. 34

```
                    195                         200                         206
W.T A.A.:           Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln

W.T. DNA:           GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA
                    CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT

pΔ222DNA:           GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA
                    CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT

A197 DNA:           GAG CTC GCA GTC ATG GCA CCT GGC GTA TCT ATC CAA
                    CTC GAG CGT CAG TAC CGT GGA CCG CAT AGA TAG GTT
                        SstI

Fragments from      GAG-CT
pΔ222 and A197      Cp
cut w/ PstI, SstI:

pΔ222, A197         GAG CTC GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA
cut & ligated       CTC GAG CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT
w/ oligodeoxy-          SstI
nucleotide pools:

                    207             210                             218
W.T A.A.:           Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn

W.T. DNA:           AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC
                    TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG

pΔ222DNA:           AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC
                    TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG

A197 DNA:           AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC
                    TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG

Fragments from      AGC ACG CTT CCC GGG AAC AAA TAC GGG GCG TAC AAC
pΔ222 and A197      TCG TGC GAA GGG CCC TTG TTT ATG CCC CGC ATG TTG
cut w/ PstI, SstI:          SmaI

                    219 220                                         230
W.T A.A.:           Gly Thr Ser Met Ala Ser Pro His Val Ala Gly Ala

W.T. DNA:           GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'
                    CCA TGC AGT TAC CGT AGA GGC GTG CAA CGG CCT CGC-5'

pΔ222DNA:           GGT ACC TCA----------CG CAC GCT GCA GGA GCG-3'
                    CCA TGG AGT----------GC GTG CGA CGT CCT CGC-5'
                     KpnI                          PstI

A197 DNA:           GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'
                    CCA TGG AGT TAC CGT AGA GGC GTG CAA GTG CCT CGC-5'

Fragments from                                              pGGA GCG-3'
pΔ222 and A197                                        A CGT CCT CGC-5'
cut w/ PstI, SstI:

pΔ222, A197         GGT ACC TCA ATG GCA TCT CCG CAC GTT GCA GGA GCG-3'
cut & ligated       CCA TGG AGT TAC CGT AGA GGC GTG CAA CGT CCT CGC-5'
w/ oligodeoxy-      KpnI                                  PstI destroyed
nucleotide pools:
```

Oligodeoxynucleotide pools synthesized with 2% contaminating nucleotides in each cycle to give ~15% of pool with 0 mutations, ~28% of pool with single mutations, and ~57% of pool with 2 or more mutations, according to the general formula $f = \frac{\mu^n}{n!} e^{-\mu}$.

FIG.—35

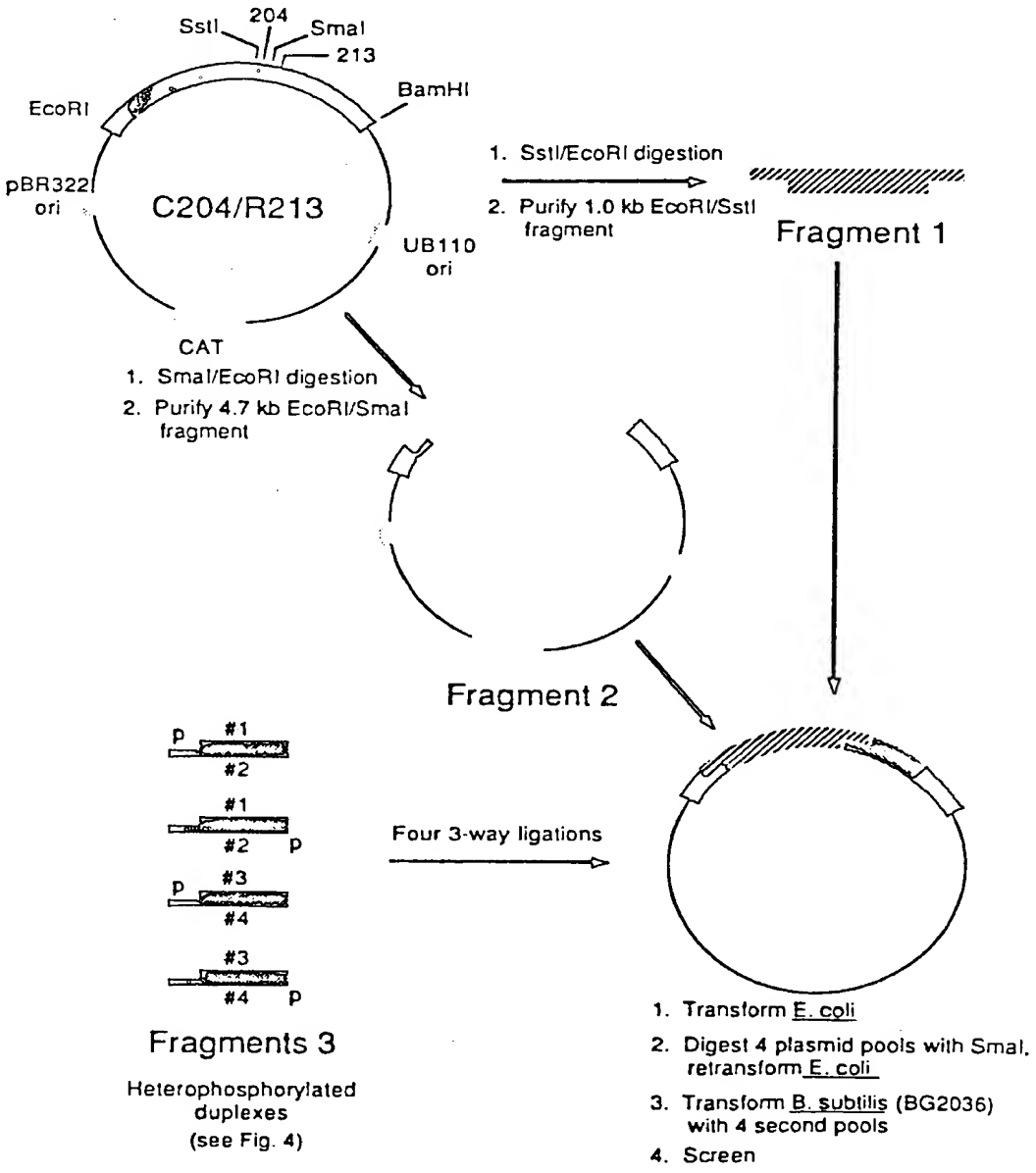

SstI
204
SmaI
213
BamHI
EcoRI
pBR322 ori
C204/R213
UB110 ori
CAT
1. SmaI/EcoRI digestion
2. Purify 4.7 kb EcoRI/SmaI fragment
1. SstI/EcoRI digestion
2. Purify 1.0 kb EcoRI/SstI fragment
Fragment 1
Fragment 2
p #1
#2
#1
#2 p
p #3
#4
#3
#4 p
Fragments 3
Heterophosphorylated duplexes
(see Fig. 4)
Four 3-way ligations
1. Transform E. coli
2. Digest 4 plasmid pools with SmaI, retransform E. coli
3. Transform B. subtilis (BG2036) with 4 second pools
4. Screen

FIG.—36

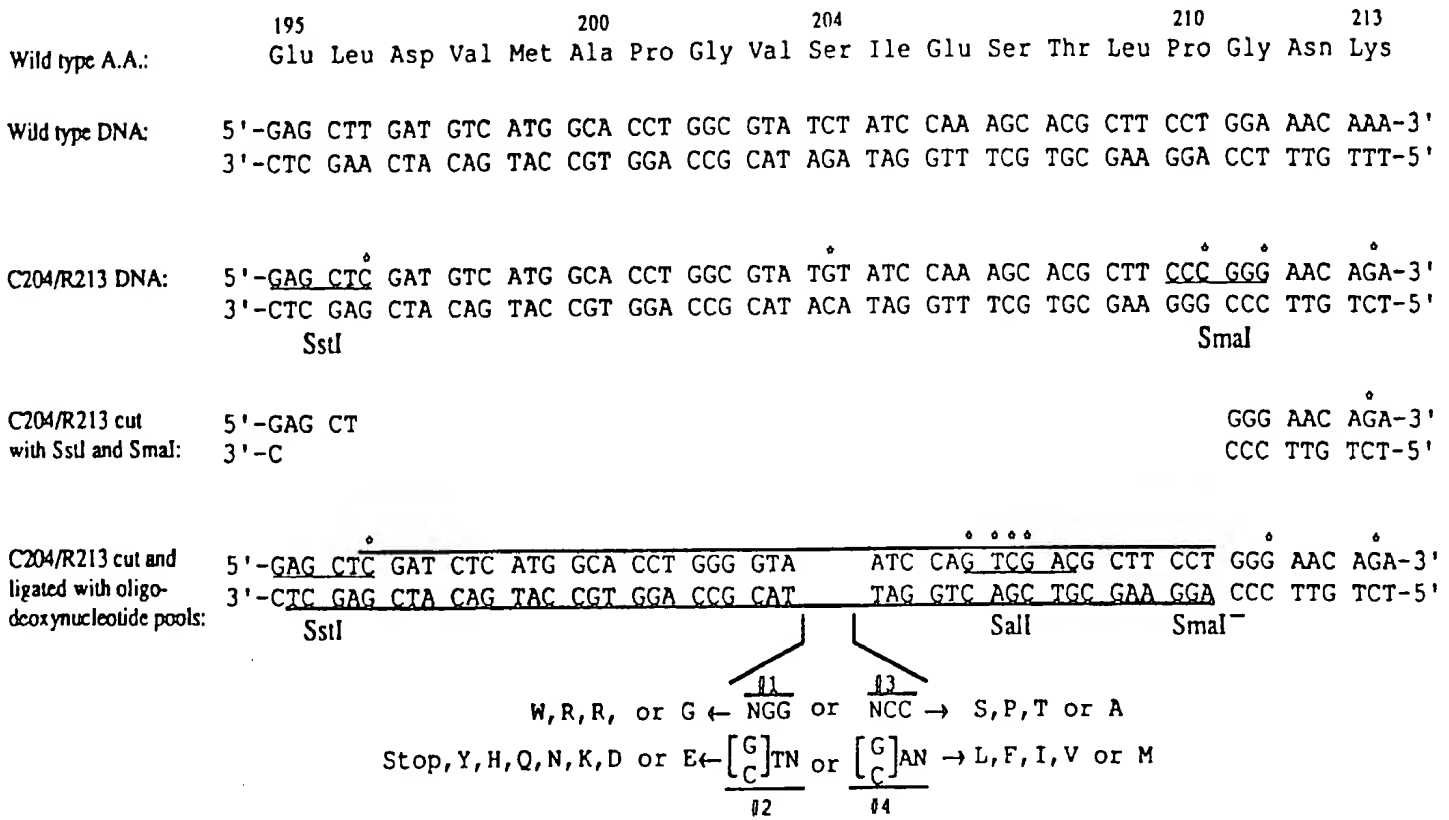

195 200 204 210 213
Wild type A.A.: Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Glu Ser Thr Leu Pro Gly Asn Lys
Wild type DNA: 5'-GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA AGC ACG CTT CCT GGA AAC AAA-3'
3'-CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT TCG TGC GAA GGA CCT TTG TTT-5'
C204/R213 DNA: 5'-GAG CTC GAT GTC ATG GCA CCT GGC GTA TGT ATC CAA AGC ACG CTT CCC GGG AAC AGA-3'
3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT ACA TAG GTT TCG TGC GAA GGG CCC TTG TCT-5'
SstI
SmaI
C204/R213 cut with SstI and SmaI: 5'-GAG CT GGG AAC AGA-3'
3'-C CCC TTG TCT-5'
C204/R213 cut and ligated with oligodeoxynucleotide pools: 5'-GAG CTC GAT CTC ATG GCA CCT GGG GTA ATC CAG TCG ACG CTT CCT GGG AAC AGA-3'
3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT TAG GTC AGC TGC GAA GGA CCC TTG TCT-5'
SstI
SalI
SmaI
#1
#3
W,R,R, or G ← NGG or NCC → S,P,T or A
Stop,Y,H,Q,N,K,D or E ← [G/C]TN or [G/C]AN → L,F,I,V or M
#2
#4

FIG.—37